

OM protein - protein search, using sw model

Run on: March 4, 2004, 15:21:50 ; Search time 104.702 Seconds
 (without alignments)
 1397.867 Million cell updates/sec

Title: US-09-668-314C-2
 Perfect score: 2687
 Sequence: 1 MGALARALLPLLAQWLLRA.....RPRDPEVVNDESSLVRHRWK 518

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
 1: geneseqp1980s:*
 2: geneseqp1990s:*
 3: geneseqp2000s:*
 4: geneseqp2001s:*
 5: genese0qp2002s:*
 6: geneseqp2003as:*
 7: geneseqp2003bs:*
 8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2687	100.0	518	2 AAW61362	Aaw61362 Aspartic
2	2687	100.0	518	2 AAY13799	Aay13799 Human asp
3	2687	100.0	518	2 AAY22239	Aay22239 Human CSP
4	2687	100.0	518	2 AAY41714	Aay41714 Human PRO
5	2687	100.0	518	3 AAY88424	Aay88424 Human asp
6	2687	100.0	518	3 AAB44270	Aab44270 Human PRO
7	2687	100.0	518	4 AAU07201	Aau07201 Human asp
8	2687	100.0	518	4 AAE10628	Aae10628 Human asp
9	2687	100.0	518	4 AAE10656	Aae10656 Human-Asp

10	2687	100.0	518	4	AAE06858	Aae06858	Human	asp
11	2687	100.0	518	4	AAE02608	Aae02608	Human	Asp
12	2687	100.0	518	4	AAE02580	Aae02580	Human	asp
13	2687	100.0	518	4	AAU29059	Aau29059	Human	PRO
14	2687	100.0	518	4	AAU06602	Aau06602	Human	Asp
15	2687	100.0	518	5	ABB06531	Abb06531	Human	asp
16	2687	100.0	518	5	ABB78589	Abb78589	Human	Asp
17	2687	100.0	518	5	ABB78617	Abb78617	Human	Asp
18	2687	100.0	518	5	ABB07453	Abb07453	Human	BAC
19	2687	100.0	518	6	ABU58435	Abu58435	Human	PRO
20	2687	100.0	518	6	ABU87983	Abu87983	Novel	hum
21	2687	100.0	518	6	ABU84298	Abu84298	Human	sec
22	2687	100.0	518	6	ABR66172	Abr66172	Human	sec
23	2687	100.0	518	6	ABR65562	Abr65562	Human	sec
24	2687	100.0	518	6	ABU99502	Abu99502	Human	sec
25	2687	100.0	518	6	ABU82741	Abu82741	Human	PRO
26	2687	100.0	518	6	ABU89862	Abu89862	Novel	hum
27	2687	100.0	518	6	ABR68111	Abr68111	Human	sec
28	2687	100.0	518	6	ABU96164	Abu96164	Novel	hum
29	2687	100.0	518	6	ABU92595	Abu92595	Human	sec
30	2687	100.0	518	6	ABO08672	Abo08672	Human	sec
31	2687	100.0	518	6	ABO02724	Abo02724	Human	sec
32	2687	100.0	518	6	ABR74878	Abr74878	Human	sec
33	2687	100.0	518	6	ABR94640	Abr94640	Human	sec
34	2687	100.0	518	6	ABO25216	Abo25216	Novel	hum
35	2687	100.0	518	6	ABU85613	Abu85613	Human	PRO
36	2687	100.0	518	6	ABU98773	Abu98773	Novel	hum
37	2687	100.0	518	6	ABU97988	Abu97988	Novel	hum
38	2687	100.0	518	6	ABU91694	Abu91694	Novel	hum
39	2687	100.0	518	6	ABU72222	Abu72222	Novel	hum
40	2687	100.0	518	6	ABU89387	Abu89387	Human	PRO
41	2687	100.0	518	6	ABU86228	Abu86228	Human	sec
42	2687	100.0	518	6	ABU67441	Abu67441	Human	sec
43	2687	100.0	518	6	ABU80469	Abu80469	Human	PRO
44	2687	100.0	518	6	ABR99387	Abr99387	Human	sec
45	2687	100.0	518	6	ABR98777	Abr98777	Human	sec

ALIGNMENTS

RESULT 1

AAW61362

ID AAW61362 standard; protein; 518 AA.

XX

AC AAW61362;

XX

DT 25-MAR-2003 (revised)

DT 25-SEP-1998 (first entry)

XX

DE Aspartic proteinase ASP1.

XX

KW ASP1; Aspartic proteinase; Alzheimer's disease; cancer; melanoma.

XX

OS Homo sapiens.

XX

PN EP848062-A2.

XX
 PD 17-JUN-1998.
 XX
 PF 01-DEC-1997; 97EP-00309648.
 XX
 PR 14-DEC-1996; 96GB-00026022.
 PR 06-OCT-1997; 97US-00999723.
 XX
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX
 PI Powell DJ, Southan C, Chapman CG, Evans JR;
 XX
 DR WPI; 1998-314477/28.
 DR N-PSDB; AAV27962.
 XX
 PT New isolated polynucleotide encodes Aspartic protease polypeptide - used
 PT to diagnosis, treat and vaccinate against Alzheimer's disease, cancer and
 PT melanoma.
 XX
 PS Claim 11; Page 7; 19pp; English.
 XX
 CC The human ASP1 protein is structurally related to other proteins of the
 CC Aspartic proteinase family. ASP 1 polypeptides and polynucleotides can be
 CC used to diagnosis, treat and vaccinate against Alzheimer's disease,
 CC cancer and melanoma. (Updated on 25-MAR-2003 to correct PR field.)
 XX
 SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 2; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.6e-231;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GS GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Db	241	GS GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360

Qy 361 YLRDENSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 |||
 Db 361 YLRDENSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

Qy 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 |||
 Db 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy 481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518
 |||
 Db 481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518

RESULT 2

AAY13799

ID AAY13799 standard; protein; 518 AA.

XX

AC AAY13799;

XX

DT 21-SEP-1999 (first entry)

XX

DE Human aspartyl protease, CSP56.

XX

KW CSP56; human; aspartyl protease; diagnosis; neoplasia; tumour;

KW breast tumour; colon tumour.

XX

OS Homo sapiens.

XX

PN WO9933963-A1.

XX

PD 08-JUL-1999.

XX

PF 14-DEC-1998; 98WO-US026547.

XX

PR 31-DEC-1997; 97US-0070112P.

XX

PA (CHIR) CHIRON CORP.

XX

PI Giese KW, Xin H;

XX

DR WPI; 1999-430240/36.

DR N-PSDB; AAX89297.

XX

PT Human CSP56 protein for diagnosis of neoplasia.

XX

PS Claim 2; Fig 2A; 51pp; English.

XX

CC This represents a human CSP56 protein, a novel aspartyl protease. The
 CC CSP56 protein can be used in methods for diagnosing neoplasia, for
 CC determining the metastatic potential of a tumour, and for screening test
 CC compounds for the ability to suppress the metastatic potential of a
 CC tumour. The tumours are preferably from breast or colon

XX

SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 2; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.6e-231;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGGPGTPAERHADGL	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GS GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Db	241	GS GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFD AVEAVARASLIPEFS DGFWTGSQLACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFD AVEAVARASLIPEFS DGFWTGSQLACWTNSETPWSYFPKISI	360
Qy	361	YLRDENS SRSFRTILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENS SRSFRTILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCA EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCA EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLVLIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLVLIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK	518

RESULT 3

AAY22239

ID AAY22239 standard; protein; 518 AA.

XX

AC AAY22239;

XX

DT 20-SEP-1999 (first entry)

XX

DE Human CSP56, aspartyl-type protease, protein sequence.

XX

KW Metastatic marker protein; human; cancer metastasis; breast cancer;
 KW colon cancer; diagnosis; therapy; tumour; metastatic potential; CSP56;
 KW aspartyl-type protease.

XX

OS Homo sapiens.

XX

PN WO9934004-A2.

XX

PD 08-JUL-1999.
 XX
 PF 24-DEC-1998; 98WO-US027608.
 XX
 PR 31-DEC-1997; 97US-0070112P.
 XX
 PA (CHIR) CHIRON CORP.
 XX
 PI Xin H, Giese K;
 XX
 DR WPI; 1999-430248/36.
 DR N-PSDB; AAX84708.
 XX
 PT New polynucleotides associated with cancer metastasis.
 XX
 PS Claim 4; Page 78-80; 80pp; English.
 XX
 CC This sequence represents a polypeptide of the invention, and is an
 CC aspartyl-type protease, designated CSP56. The polynucleotides (PNs) of
 CC the invention encode metastatic marker protein variants. The PNs and
 CC polypeptides can be used as markers for cancer metastasis. The products
 CC can be used for identifying metastatic tissue or metastatic potential of
 CC a tissue, e.g. breast or colon tissue. They can also be used for
 CC screening test compounds for the ability to suppress the metastatic
 CC potential of a tumour. The products can be used for developing products
 CC for the therapy of cancers, particularly breast or colon cancer
 XX
 SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 2; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.6e-231;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
Qy	61	ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360

Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 |||
 Db 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

Qy 421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 |||
 Db 421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy 481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518
 |||
 Db 481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518

RESULT 4

AAY41714

ID AAY41714 standard; protein; 518 AA.

XX

AC AAY41714;

XX

DT 07-DEC-1999 (first entry)

XX

DE Human PRO852 protein sequence.

XX

KW Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
 KW probe; blood coagulation disorder; cancer; cellular adhesion disorder;
 KW secreted protein; transmembrane protein.

XX

OS Homo sapiens.

XX

PN WO9946281-A2.

XX

PD 16-SEP-1999.

XX

PF 08-MAR-1999; 99WO-US005028.

XX

PR 10-MAR-1998; 98US-0077450P.

PR 11-MAR-1998; 98US-0077632P.

PR 11-MAR-1998; 98US-0077641P.

PR 11-MAR-1998; 98US-0077649P.

PR 12-MAR-1998; 98US-0077791P.

PR 13-MAR-1998; 98US-0078004P.

PR 17-MAR-1998; 98US-00040220.

PR 20-MAR-1998; 98US-0078886P.

PR 20-MAR-1998; 98US-0078910P.

PR 20-MAR-1998; 98US-0078936P.

PR 20-MAR-1998; 98US-0078939P.

PR 25-MAR-1998; 98US-0079294P.

PR 26-MAR-1998; 98US-0079656P.

PR 27-MAR-1998; 98US-0079663P.

PR 27-MAR-1998; 98US-0079664P.

PR 27-MAR-1998; 98US-0079689P.

PR 27-MAR-1998; 98US-0079728P.

PR 27-MAR-1998; 98US-0079786P.

PR 30-MAR-1998; 98US-0079920P.

PR 30-MAR-1998; 98US-0079923P.

PR 31-MAR-1998; 98US-0080105P.

PR 31-MAR-1998; 98US-0080107P.

PR	31-MAR-1998;	98US-0080165P.
PR	31-MAR-1998;	98US-0080194P.
PR	01-APR-1998;	98US-0080327P.
PR	01-APR-1998;	98US-0080328P.
PR	01-APR-1998;	98US-0080333P.
PR	01-APR-1998;	98US-0080334P.
PR	08-APR-1998;	98US-0081049P.
PR	08-APR-1998;	98US-0081070P.
PR	08-APR-1998;	98US-0081071P.
PR	09-APR-1998;	98US-0081195P.
PR	09-APR-1998;	98US-0081203P.
PR	09-APR-1998;	98US-0081229P.
PR	15-APR-1998;	98US-0081817P.
PR	15-APR-1998;	98US-0081838P.
PR	15-APR-1998;	98US-0081952P.
PR	15-APR-1998;	98US-0081955P.
PR	21-APR-1998;	98US-0082568P.
PR	21-APR-1998;	98US-0082569P.
PR	22-APR-1998;	98US-0082700P.
PR	22-APR-1998;	98US-0082704P.
PR	22-APR-1998;	98US-0082804P.
PR	23-APR-1998;	98US-0082767P.
PR	23-APR-1998;	98US-0082796P.
PR	27-APR-1998;	98US-0083336P.
PR	28-APR-1998;	98US-0083322P.
PR	29-APR-1998;	98US-0083392P.
PR	29-APR-1998;	98US-0083495P.
PR	29-APR-1998;	98US-0083496P.
PR	29-APR-1998;	98US-0083499P.
PR	29-APR-1998;	98US-0083500P.
PR	29-APR-1998;	98US-0083545P.
PR	29-APR-1998;	98US-0083554P.
PR	29-APR-1998;	98US-0083558P.
PR	29-APR-1998;	98US-0083559P.
PR	30-APR-1998;	98US-0083742P.
PR	05-MAY-1998;	98US-0084366P.
PR	06-MAY-1998;	98US-0084414P.
PR	06-MAY-1998;	98US-0084441P.
PR	07-MAY-1998;	98US-0084598P.
PR	07-MAY-1998;	98US-0084600P.
PR	07-MAY-1998;	98US-0084627P.
PR	07-MAY-1998;	98US-0084637P.
PR	07-MAY-1998;	98US-0084639P.
PR	07-MAY-1998;	98US-0084640P.
PR	07-MAY-1998;	98US-0084643P.
PR	13-MAY-1998;	98US-0085323P.
PR	13-MAY-1998;	98US-0085338P.
PR	13-MAY-1998;	98US-0085339P.
PR	15-MAY-1998;	98US-0085573P.
PR	15-MAY-1998;	98US-0085579P.
PR	15-MAY-1998;	98US-0085580P.
PR	15-MAY-1998;	98US-0085582P.
PR	15-MAY-1998;	98US-0085689P.
PR	15-MAY-1998;	98US-0085697P.
PR	15-MAY-1998;	98US-0085700P.
PR	15-MAY-1998;	98US-0085704P.
PR	18-MAY-1998;	98US-0086023P.

PR	22-MAY-1998;	98US-0086392P.
PR	22-MAY-1998;	98US-0086414P.
PR	22-MAY-1998;	98US-0086430P.
PR	22-MAY-1998;	98US-0086486P.
PR	28-MAY-1998;	98US-0087098P.
PR	28-MAY-1998;	98US-0087106P.
PR	28-MAY-1998;	98US-0087208P.
PR	30-JUL-1998;	98US-0094651P.
PR	11-SEP-1998;	98US-0100038P.

PA (GETH) GENENTECH INC.

PI Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;

DR WPI; 1999-551358/46.

PT New secreted and transmembrane polypeptides and their polynucleotides,
PT useful for treating blood coagulation disorders, cancers and cellular
PT adhesion disorders.

PS Claim 12; Fig 73; 530pp; English.

CC The present invention describes secreted and transmembrane polypeptides
CC and their polynucleotides. The nucleotide sequences are useful as sources
CC of probes, primers, for chromosome mapping, and for generation of
CC antisense sequences. They can also be used to create transgenic animals.
CC The proteins can be used to treat a variety of diseases and disorders,
CC depending on their function. Diseases that may be treated include blood
CC coagulation disorders, cancers and cellular adhesion disorders. They may
CC also be used to raise antibodies. AAZ33891 to AAZ34338, and AAY41685 to
CC AAY41774 represent polynucleotide and polypeptide sequence given in the
CC exemplification of the present invention

S0 Sequence 518 AA;

```
Query Match      100.0%;   Score 2687;   DB 2;   Length 518;
Best Local Similarity 100.0%;   Pred. No. 8.6e-231;
Matches 518;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
```

Qy 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGGPGTPAERHADGL 60
 |||

Db 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGGPGTPAERHADGL 60

Qy 61 ALALEPALAS PAGAANFLAMVDNLQGD SGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
 |||

Db 61 ALALEPALAS PAGAANFLAMVDNLQGD SGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy	121	TPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180

Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300

Db	241	GS GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDS GTTLLRLPQKVFD AVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Db	301	IVDS GTTLLRLPQKVFD AVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Qy	361	YLRDENS SRSFRTILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENS SRSFRTILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGF AASPCA EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGF AASPCA EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLV LIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLV LIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK	518

RESULT 5

AA Y88424

ID AAY88424 standard; protein; 518 AA.

XX

AC AAY88424;

XX

DT 03-AUG-2000 (first entry)

XX

DE Human aspartyl protease 1 (Asp1) amino acid sequence.

XX

KW Aspartyl protease; aspartase; amyloid precursor protein; APP; Asp 1;

KW Alzheimer's disease; beta secretase site.

XX

OS Homo sapiens.

XX

PN WO200017369-A2.

XX

PD 30-MAR-2000.

XX

PF 23-SEP-1999; 99WO-US020881.

XX

PR 24-SEP-1998; 98US-0101594P.

XX

PA (PHAA) PHARMACIA & UPJOHN CO.

XX

PI Gurney ME, Bienkowski MJ, Heinrikson RL, Parodi LA, Yan R;

XX

DR WPI; 2000-303209/26.

DR

N-PSDB; AAA15661.

XX

PT New enzyme designated human aspartase useful in research into Alzheimer's

PT

Disease is capable of cleaving amyloid protein precursor at the beta

PT

secretase site to produce amyloid beta peptide.

XX

PS Claim 54; Fig 1; 183pp; English.

XX

CC This sequence represents the human aspartyl protease amino acid sequence.

CC The invention relates to a protease capable of cleaving the beta
CC secretase site of amyloid precursor protein (APP). The protease contains
CC a sequence encoding the amino acid sequence DTG and a sequence encoding
CC DSG or DTG separated by 100-300 amino acids. When mutated the APP gene
CC causes an autosomal dominant form of Alzheimer's disease. APP localises
CC to the cell surface membrane and have a single C-terminal transmembrane
CC domain. Proteolytic processing of APP produces the amyloid beta protein,
CC which is possibly very important in Alzheimer's disease. The invention
CC includes a nucleotide sequence encoding the protease, a vector containing
CC the nucleotide sequence, and a cell line comprising the vector. Methods
CC for screening for inhibitors of beta secretase activity are also given in
CC the invention. The human aspartase protein and nucleotide sequences and
CC the methods for identifying inhibitors of the protease, are useful in the
CC treatment of and research in to Alzheimer's disease

XX

SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 3; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.6e-231;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
|||||

Db 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy 61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
|||||

Db 61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy 121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
|||||

Db 121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180

Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
|||||

Db 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
|||||

Db 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300

Qy 301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISI 360
|||||

Db 301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISI 360

Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
|||||

Db 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

Qy 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
|||||

Db 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518
|||||

Db 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518

RESULT 6

AAB44270

ID AAB44270 standard; protein; 518 AA.

XX

AC AAB44270;

XX

DT 08-FEB-2001 (first entry)

XX

DE Human PRO852 (UNQ418) protein sequence SEQ ID NO:196.

XX

KW Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;
KW expressed sequence tag; detection; cancer.

XX

OS Homo sapiens.

XX

PN WO200053756-A2.

XX

PD 14-SEP-2000.

XX

PF 18-FEB-2000; 2000WO-US004341.

XX

PR 08-MAR-1999; 99WO-US005028.

PR 12-MAR-1999; 99US-0123957P.

PR 29-MAR-1999; 99US-0126773P.

PR 21-APR-1999; 99US-0130232P.

PR 28-APR-1999; 99US-0131445P.

PR 14-MAY-1999; 99US-0134287P.

PR 23-JUN-1999; 99US-0141037P.

PR 26-JUL-1999; 99US-0145698P.

PR 29-OCT-1999; 99US-0162506P.

PR 30-NOV-1999; 99WO-US028313.

PR 02-DEC-1999; 99WO-US028551.

PR 02-DEC-1999; 99WO-US028565.

PR 16-DEC-1999; 99WO-US030095.

PR 30-DEC-1999; 99WO-US031243.

PR 30-DEC-1999; 99WO-US031274.

PR 05-JAN-2000; 2000WO-US000219.

PR 06-JAN-2000; 2000WO-US000277.

PR 06-JAN-2000; 2000WO-US000376.

XX

PA (GETH) GENENTECH INC.

XX

PI Ashkenazi AJ, Baker KP, Botstein D, Desnoyers L, Eaton DL;

PI Ferrara N, Filvaroff E, Fong S, Gao W, Gerber H, Gerritsen ME;

PI Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL, Hillan KJ;

PI Kljavin IJ, Kuo SS, Napier MA, Pan J, Paoni NF, Roy MA, Shelton DL;

PI Stewart TA, Tumas D, Williams PM, Wood WI;

XX

DR WPI; 2000-611443/58.

DR N-PSDB; AAC78500.

XX

PT Novel PRO polypeptides and polynucleotides used in detection methods, to
PT target bioactive molecules to specific cells, and to modulate cellular
PT activities.

XX

PS Claim 12; Fig 73; 636pp; English.

AAU07201

ID AAU07201 standard; protein; 518 AA.

XX

AC AAU07201;

XX

DT 24-OCT-2001 (first entry)

XX

DE Human aspartyl protease 1 (Asp-1).

XX

KW Human; aspartyl protease 1; Asp-1; nootropic; neuroprotective;

KW aspartyl protease 2; Asp2; amyloid protein precursor; APP;

KW beta-secretase; Alzheimer's disease.

XX

OS Homo sapiens.

XX

PN WO200149097-A2.

XX

PD 12-JUL-2001.

XX

PF 09-MAY-2001; 2001WO-IB000797.

XX

PR 09-MAY-2001; 2001WO-IB000797.

XX

PA (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

PA (YANR/) YAN R.

XX

PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;

XX

DR WPI; 2001-502548/55.

DR N-PSDB; AAS11701.

XX

PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity.

XX

PS Example 2; Fig 1; 185pp; English.

XX

CC The invention relates to a novel purified polypeptide comprising a
CC fragment of mammalian aspartyl protease 2 (Asp2) protein which lacks the
CC Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide
CC and the fragment retain the beta-secretase activity of the mammalian Asp2
CC protein. Also included is an isoform of amyloid protein precursor (APP)
CC comprising the amino acid sequence of a APP or its fragment containing an
CC APP cleavage site recognisable by a mammalian beta-secretase, and further
CC comprising two lysine residues at the carboxyl terminus of the amino acid
CC sequence of the mammalian APP or APP fragment. The polypeptides are used
CC for assaying for modulators of beta-secretase activity; identifying
CC agents that inhibit the APP processing activity of human Asp2 aspartyl
CC protease (Hu-Asp2); identifying agents that modulate the activity of Asp2
CC ; and for reducing cellular production of amyloid beta (Abeta) from APP.
CC Agents identified by the above methods are useful for treating
CC Alzheimer's disease; and for identifying modulators of amyloid-beta
CC (Abeta) peptide production, for use in designing therapeutics for the

CC treatment or prevention of Alzheimer's disease. Probes and primers
 CC derived from Asp nucleic acid sequences are useful for detecting Hu-Asp
 CC nucleic acids in in vitro assays and in Northern and Southern blots. The
 CC present sequence represents the amino acid sequence of human Asp-1
 XX
 SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 4; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.6e-231;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
Qy	61	ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLVLIIVLVLPLPFRQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLVLIIVLVLPLPFRQRRPRDPEVVNDESSLVRHRWK	518

RESULT 8

AAE10628

ID AAE10628 standard; protein; 518 AA.

XX

AC AAE10628;

XX

DT 10-DEC-2001 (first entry)

XX

DE Human aspartyl protease 1 (hu-Asp1) protein.

XX
 KW Human; aspartyl protease 1; Aspl; amyloid precursor protein; APP;
 KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
 KW amyloid plaque; neuronal loss; proteolytic; nootropic; neuroprotective;
 KW chromosome 21.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1. .20
 FT /label= Signal_peptide
 FT Protein 21. .518
 FT /note= "Mature human aspartyl protease 1"
 FT Domain 469. .492
 FT /label= Transmembrane_domain
 XX
 PN GB2357767-A.
 XX
 PD 04-JUL-2001.
 XX
 PF 22-SEP-2000; 2000GB-00023315.
 XX
 PR 23-SEP-1999; 99US-00404133.
 PR 23-SEP-1999; 99US-0155493P.
 PR 23-SEP-1999; 99WO-US020881.
 PR 13-OCT-1999; 99US-00416901.
 PR 06-DEC-1999; 99US-0169232P.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Bienkowski MJ, Gurney M;
 XX
 DR WPI; 2001-444208/48.
 DR N-PSDB; AAD17864.
 XX
 PT Polypeptide comprising fragments of human aspartyl protease with amyloid
 PT precursor protein processing activity and alpha-secretase activity, for
 PT identifying modulators useful in treating Alzheimer's disease.
 XX
 PS Claim 36; Fig 1; 187pp; English.
 XX
 CC The patent discloses human aspartyl protease 1 (hu-Aspl) or modified Aspl
 CC proteins which lack transmembrane domain or amino terminal domain or
 CC cytoplasmic domain and retains alpha-secretase activity and amyloid
 CC protein precursor (APP) processing activity. The proteins of the
 CC invention are useful for assaying hu-Aspl alpha-secretase activity, which
 CC in turn is useful for identifying modulators of hu-Aspl alpha-secretase
 CC activity, where modulators that increase hu-Aspl alpha-secretase activity
 CC are useful for treating Alzheimer's disease (AD) which causes progressive
 CC dementia with consequent formation of amyloid plaques, neurofibrillary
 CC tangles, gliosis and neuronal loss. Hu-Aspl protease substrate is useful
 CC for assaying hu-Aspl proteolytic activity, by contacting hu-Aspl protein
 CC with the substrate under acidic conditions and determining the level of
 CC hu-Aspl proteolytic activity. The present sequence is Aspl protein from
 CC human. Aspl gene is localised on chromosome 21
 XX
 SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.6e-231;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy     61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGENTSFLVNIATI 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGENTSFLVNIATI 180

Qy    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA 300

Qy    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360

Qy    361 YLRDENSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 YLRDENSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

Qy    421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy    481 AILLVLIVLLLLLPPFRQRRPRDPEVVNDESSLVRHRWK 518
        ||||||||||||||||||||||||||||||||||||||||
Db    481 AILLVLIVLLLLLPPFRQRRPRDPEVVNDESSLVRHRWK 518
```

RESULT 9

AAE10656

ID AAE10656 standard; protein; 518 AA.

XX

AC AAE10656;

XX

DT 10-DEC-2001 (first entry)

XX

DE Human-Asp 1 protein lacking TM domain and containing (His)6 tag.

XX

KW Human; aspartyl protease 1; Aspl; amyloid precursor protein; APP;

KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;

KW amyloid plaque; neuronal loss; proteolytic; nootropic; neuroprotective.

XX

OS Homo sapiens.

OS Synthetic.
 XX
 PN GB2357767-A.
 XX
 PD 04-JUL-2001.
 XX
 PF 22-SEP-2000; 2000GB-00023315.
 XX
 PR 23-SEP-1999; 99US-00404133.
 PR 23-SEP-1999; 99US-0155493P.
 PR 23-SEP-1999; 99WO-US020881.
 PR 13-OCT-1999; 99US-00416901.
 PR 06-DEC-1999; 99US-0169232P.
 XX
 PA (PHAA) PHARMACIA & UPJOHN CO.
 XX
 PI Bienkowski MJ, Gurney M;
 XX
 DR WPI; 2001-444208/48.
 XX
 PT Polypeptide comprising fragments of human aspartyl protease with amyloid
 PT precursor protein processing activity and alpha-secretase activity, for
 PT identifying modulators useful in treating Alzheimer's disease.
 XX
 PS Example 14; Page 155-156; 187pp; English.
 XX
 CC The patent discloses human aspartyl protease 1 (hu-Asp1) or modified Asp1
 CC proteins which lack transmembrane domain or amino terminal domain or
 CC cytoplasmic domain and retains alpha-secretase activity and amyloid
 CC protein precursor (APP) processing activity. The proteins of the
 CC invention are useful for assaying hu-Asp1 alpha-secretase activity, which
 CC in turn is useful for identifying modulators of hu-Asp1 alpha-secretase
 CC activity, where modulators that increase hu-Asp1 alpha-secretase activity
 CC are useful for treating Alzheimer's disease (AD) which causes progressive
 CC dementia with consequent formation of amyloid plaques, neurofibrillary
 CC tangles, gliosis and neuronal loss. Hu-Asp1 protease substrate is useful
 CC for assaying hu-Asp1 proteolytic activity, by contacting hu-Asp1 protein
 CC with the substrate under acidic conditions and determining the level of
 CC hu-Asp1 proteolytic activity. The present sequence is human Asp 1 protein
 CC lacking a transmembrane (TM) domain and containing (His)6 tag. This
 CC sequence is generated from human Asp 1 protein by the deletion of its C-
 CC terminal TM domain and addition of hexa-histidine tag at its C-terminus
 XX
 SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 4; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.6e-231;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

 Qy 61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGEVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGEVGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GS GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Db	241	GS GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAEIAGA AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAEIAGA AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLVLIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLVLIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK	518

RESULT 10

AAE06858

ID AAE06858 standard; protein; 518 AA.

XX

AC AAE06858;

XX

DT 23-OCT-2001 (first entry)

XX

DE Human aspartyl protease 1 (Hu-Aspl) protein.

XX

KW Human; aspartyl protease 1; Asp 1; beta-amyloid precursor protein; APP;

KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;

KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; nootropic;

KW neuroprotective; antisense therapy; gene therapy; chromosome 21.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	Peptide	1. .20
----	---------	--------

FT		/label= Signal_peptide
----	--	------------------------

FT	Protein	21. .518
----	---------	----------

FT		/note= "Mature human aspartyl protease 1 (Hu-Aspl)"
----	--	---

FT	Domain	469. .492
----	--------	-----------

FT		/label= Transmembrane_domain
----	--	------------------------------

XX

PN WO200150829-A2.

XX

PD 19-JUL-2001.

XX
PF 09-MAY-2001; 2001WO-IB000799.
XX
PR 09-MAY-2001; 2001WO-IB000799.
XX
PA (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX
PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX
DR WPI; 2001-483072/52.
DR N-PSDB; AAD13020.
XX
PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity.
XX
PS Example 2; Fig 1; 185pp; English.
XX
CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
CC precursor protein (APP) isoforms and their corresponding DNA molecules.
CC Human aspartyl proteases can act as beta-secretase proteases useful for
CC treating Alzheimer's disease. APP isoforms are useful for identifying
CC modulators of amyloid-beta peptide production, for use in designing
CC therapeutics for the treatment and prevention of Alzheimer's disease,
CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
CC and neuronal loss. APP isoforms are also used in methods for identifying
CC inhibitors and modulators of human Asp2 activity. The invention relates
CC to a method for identifying agents that modulate the activity of human
CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
CC as a means to screen in cellular assays for the inhibitors of beta- and
CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
CC polymerase chain reactions (PCR). The probes are useful for detecting Hu-
CC Asp nucleic acids in in vitro assays and in Northern and Southern blots.
CC The present sequence is human aspartyl protease 1 (Hu-Asp1). Hu-Asp 1
CC gene is localised on chromosome 21
XX
SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.6e-231;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
|
Db 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy 61 ALALEPALASPAGAANFLAMVDNLQDSDGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
|
Db 61 ALALEPALASPAGAANFLAMVDNLQDSDGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy 121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI 180
|

Db 121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGEVGEDLVTIPKGFNTSFLVNIATI 180
 Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 Qy 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 Qy 301 IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 Qy 421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 Qy 481 AILLVLIVLLLLPFCRQRRPRDPEVVNDESSLVRHRWK 518
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 481 AILLVLIVLLLLPFCRQRRPRDPEVVNDESSLVRHRWK 518

RESULT 11

AAE02608

ID AAE02608 standard; protein; 518 AA.

XX

AC AAE02608;

XX

DT 10-AUG-2001 (first entry)

XX

DE Human Aspartyl protease-1 (Asp-1) deltaTM (His)6 protein.

XX

KW Human; alpha-secretase; amyloid precursor protein; APP; therapy;

KW Alzheimer's disease; antialzheimer's; aspartyl protease 1; Aspl;

KW beta-secretase; Asp-1 deltaTM (His)6 protein.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200123533-A2.

XX

PD 05-APR-2001.

XX

PF 22-SEP-2000; 2000WO-US026080.

XX

PR 23-SEP-1999; 99US-0155493P.

PR 23-SEP-1999; 99WO-US020881.

PR 13-OCT-1999; 99US-00416901.

PR 06-DEC-1999; 99US-0169232P.

XX

PA (PHAA) PHARMACIA & UPJOHN CO.

XX

PI Gurney M, Bienkowski MJ;
XX
DR WPI; 2001-290516/30.
XX

PT Enzymes that cleave the alpha-secretase site of the amyloid precursor
PT protein, useful for the treatment of Alzheimer's disease.
XX

PS Example 14; Page 183-184; 189pp; English.
XX

CC The present invention relates to enzymes for cleaving the alpha-
CC secretase site of the amyloid precursor protein (APP) and methods of
CC identifying those enzymes. The methods may be used to identify enzymes
CC that may be used to cleave the alpha-secretase cleavage site of the APP
CC protein. The enzymes may be used to treat or modulate the progress of
CC Alzheimer's disease. The present sequence is human Aspartyl protease-1
CC (Asp-1) deltaTM (His)6 protein which is used for the expression of pre-
CC pro-human-Aspartyl protease 1 (Asp1). This protein is obtained by
CC replacing C-terminal transmembrane and cytoplasmic domains with a
CC hexahistidine purification tag in the human Aspartyl protease 1
XX

SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 4; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.6e-231;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAANFLAMVDNLQDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAANFLAMVDNLQDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFD AVEAVARASLIPEFS DGFWTGSQLACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFD AVEAVARASLIPEFS DGFWTGSQLACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480

QY 481 AILLVLIVLLLLPFCRQRRPRDPEVVNDESSLVRHRWK 518
 |||||
 Db 481 AILLVLIVLLLLPFCRQRRPRDPEVVNDESSLVRHRWK 518

RESULT 12

AAE02580

ID AAE02580 standard; protein; 518 AA.

XX

AC AAE02580;

XX

DT 10-AUG-2001 (first entry)

XX

DE Human aspartyl protease 1 (Asp 1).

XX

KW Human; alpha-secretase; amyloid precursor protein; APP; therapy;

KW Alzheimer's disease; antialzheimer's; aspartyl protease 1; Asp 1;

KW beta-secretase; chromosome 21.

XX

OS Homo sapiens.

XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	Peptide	1. .20
----	---------	--------

FT		/label= Signal_peptide
----	--	------------------------

FT	Peptide	22. .62
----	---------	---------

FT		/label= Asp_1_prepropeptide
----	--	-----------------------------

FT	Peptide	23. .62
----	---------	---------

FT		/label= Asp_1_propeptide
----	--	--------------------------

FT	Protein	63. .518
----	---------	----------

FT		/label= Mature_human_Asp_1_protein
----	--	------------------------------------

FT		/note= "Specifically claimed"
----	--	-------------------------------

FT	Active-site	87. .89
----	-------------	---------

FT		/label= Active_site_1
----	--	-----------------------

FT	Active-site	110. .113
----	-------------	-----------

FT		/label= Active_site_2
----	--	-----------------------

FT	Active-site	303. .305
----	-------------	-----------

FT		/label= Active_site_3
----	--	-----------------------

FT	Domain	469. .492
----	--------	-----------

FT		/label= Transmembrane_domain
----	--	------------------------------

FT	Domain	493. .518
----	--------	-----------

FT		/label= Cytoplasmic_domain
----	--	----------------------------

FT	Region	497. .518
----	--------	-----------

FT		/note= "Peptide #1"
----	--	---------------------

XX

PN WO200123533-A2.

XX

PD 05-APR-2001.

XX

PF 22-SEP-2000; 2000WO-US026080.

XX

PR 23-SEP-1999; 99US-0155493P.

PR 23-SEP-1999; 99WO-US020881.

PR 13-OCT-1999; 99US-00416901.

PR 06-DEC-1999; 99US-0169232P.

XX

PA (PHAA) PHARMACIA & UPJOHN CO.

XX
 PI Gurney M, Bienkowski MJ;
 XX
 DR WPI; 2001-290516/30.
 DR N-PSDB; AAD06738.
 XX
 PT Enzymes that cleave the alpha-secretase site of the amyloid precursor
 PT protein, useful for the treatment of Alzheimer's disease.
 XX
 PS Claim 29; Fig 1; 189pp; English.
 XX
 CC The present invention relates to enzymes for cleaving the alpha-
 CC secretase site of the amyloid precursor protein (APP) and methods of
 CC identifying those enzymes. The methods may be used to identify enzymes
 CC that may be used to cleave the alpha-secretase cleavage site of the APP
 CC protein. The enzymes may be used to treat or modulate the progress of
 CC Alzheimer's disease. The present sequence is human aspartyl protease 1
 CC (Asp 1). Asp 1 has alpha-secretase protease and beta-secretase protease
 CC activities. Asp 1 gene is located on chromosome 21
 XX
 SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 4; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.6e-231;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
Db	1	MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
Qy	61	ALALEPALASPAGAANFLAMVDNLQGD	120
Db	61	ALALEPALASPAGAANFLAMVDNLQGD	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDVT	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDVT	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Db	241	GGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFD	360
Db	301	IVDSGTTLLRLPQKVFD	360
Qy	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480

QY 481 AILLVLIVLLLLPFCQRRPRDPEVVNDESSLVRHRWK 518
 |||||
 Db 481 AILLVLIVLLLLPFCQRRPRDPEVVNDESSLVRHRWK 518

RESULT 13

AAU29059

ID AAU29059 standard; protein; 518 AA.

XX

AC AAU29059;

XX

DT 18-DEC-2001 (first entry)

XX

DE Human PRO polypeptide sequence #36.

XX

KW PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;

KW dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;

KW blood; chondrocyte cell; cell proliferation; cell differentiation; colon;

KW adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.

XX

OS Homo sapiens.

XX

PN WO200168848-A2.

XX

PD 20-SEP-2001.

XX

PF 28-FEB-2001; 2001WO-US006520.

XX

PR 01-MAR-2000; 2000WO-US005601.

PR 02-MAR-2000; 2000WO-US005841.

PR 03-MAR-2000; 2000US-0187202P.

PR 06-MAR-2000; 2000US-0186968P.

PR 14-MAR-2000; 2000US-0189320P.

PR 14-MAR-2000; 2000US-0189328P.

PR 15-MAR-2000; 2000WO-US006884.

PR 21-MAR-2000; 2000US-0190828P.

PR 21-MAR-2000; 2000US-0191007P.

PR 21-MAR-2000; 2000US-0191048P.

PR 21-MAR-2000; 2000US-0191314P.

PR 28-MAR-2000; 2000US-0192655P.

PR 29-MAR-2000; 2000US-0193032P.

PR 29-MAR-2000; 2000US-0193053P.

PR 30-MAR-2000; 2000WO-US008439.

PR 04-APR-2000; 2000US-0194449P.

PR 04-APR-2000; 2000US-0194647P.

PR 11-APR-2000; 2000US-0195975P.

PR 11-APR-2000; 2000US-0196000P.

PR 11-APR-2000; 2000US-0196187P.

PR 11-APR-2000; 2000US-0196690P.

PR 11-APR-2000; 2000US-0196820P.

PR 18-APR-2000; 2000US-0198121P.

PR 18-APR-2000; 2000US-0198585P.

PR 25-APR-2000; 2000US-0199397P.

PR 25-APR-2000; 2000US-0199550P.

PR 25-APR-2000; 2000US-0199654P.

PR 03-MAY-2000; 2000US-0201516P.

Qy	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGLIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGLIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLVLIIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLVLIIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK	518

RESULT 14

AAU06602

ID AAU06602 standard; protein; 518 AA.

XX

AC AAU06602;

XX

DT 24-OCT-2001 (first entry)

XX

DE Human Aspartyl protease 1 (Asp1).

XX

KW Human; Aspartyl protease; Asp1; Asp2; beta-secretase; nootropic;
 KW neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
 KW amyloid-beta; Abeta.

XX

OS Homo sapiens.

XX

PN WO200149098-A2.

XX

PD 12-JUL-2001.

XX

PF 09-MAY-2001; 2001WO-IB000798.

XX

PR 09-MAY-2001; 2001WO-IB000798.

XX

PA (BIEN/) BIENKOWSKI M J.

PA (GURN/) GURNEY M E.

PA (HEIN/) HEINRIKSON R L.

PA (PARO/) PARODI L A.

PA (YANR/) YAN R.
 XX
 PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
 XX
 DR WPI; 2001-502549/55.
 DR N-PSDB; AAS11516.
 XX
 PT Novel purified polypeptide comprising fragment of mammalian aspartyl
 PT protease 2, lacking Asp2 transmembrane domain and retaining beta
 PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
 PT activity.
 XX
 PS Example 2; Fig 1; 185pp; English.
 XX
 CC The invention relates to a purified polypeptide comprising a fragment of
 CC mammalian aspartyl protease (Asp)2 protein which lacks the Asp2
 CC transmembrane domain and the Asp2 protein, and where the polypeptide and
 CC the fragment retain the beta-secretase activity of the mammalian Asp2
 CC protein. The invention also details polynucleotides for the Asp proteins
 CC and vectors expressing them, and a polypeptide (isoform of amyloid
 CC protein precursor (APP)) comprising the amino acid sequence of an APP or
 CC its fragment containing an APP cleavage site recognizable by a mammalian
 CC beta-secretase, and further comprising two lysine residues at the
 CC carboxyl terminus of the amino acid sequence of the mammalian APP or APP
 CC fragment. Also included in the invention are methods of identifying
 CC modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
 CC useful for treating Alzheimer's disease. APP is useful in methods for
 CC identifying inhibitors or modulators of human Asp2 activity and amyloid-
 CC beta (Abeta) peptide production. APP is also useful in designing
 CC therapeutics for the treatment or prevention of Alzheimer's disease. APP
 CC comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which is
 CC associated with increased levels of Abeta processing is useful in assays
 CC relating the Alzheimer's research. The expression vector is useful for
 CC recombinantly expressing APP. Nucleic acids that hybridise to Asp
 CC oligonucleotides are useful as probes or primers. The probes are useful
 CC for detecting Hu-Asp nucleic acids in in vitro assays and in Northern and
 CC Southern blots. The present sequence is human Asp1
 XX
 SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 4; Length 518;
 Best Local Similarity 100.0%; Pred. No. 8.6e-231;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGGPGTPAERHADGL 60
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGGPGTPAERHADGL 60
 Qy 61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
 Qy 121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
 Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAEIAGA AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAEIAGA AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLV LIVLLLLPFCQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLV LIVLLLLPFCQRRPRDPEVVNDESSLVRHRWK	518

RESULT 15

ABB06531

ID ABB06531 standard; protein; 518 AA.

XX

AC ABB06531;

XX

DT 31-MAY-2002 (first entry)

XX

DE Human aspartyl protease 1 protein SEQ ID NO:125.

XX

KW Beta-secretase; enzyme; cleavage site; amyloid protein precursor; APP;
 KW aspartyl protease; neuroprotective; nootropic; beta-secretase inhibitor;
 KW Alzheimer's disease.

XX

OS Homo sapiens.

XX

PN WO200206306-A2.

XX

PD 24-JAN-2002.

XX

PF 19-JUL-2001; 2001WO-US023035.

XX

PR 19-JUL-2000; 2000US-0219795P.

PR 12-MAR-2001; 2001US-0275251P.

XX

PA (PHAA) PHARMACIA & UPJOHN CO.

XX

PI Yan R, Tomasselli AG, Gurney ME, Emmons TL, Bienkowski MJ;
 PI Heinrikson RL;

XX

DR WPI; 2002-216995/27.

XX

PT Novel substrates for human aspartyl protease useful for identifying

PT modulators of beta secretase activity of aspartyl protease for treating
PT Alzheimer's disease.

XX

PS Disclosure; Page 161-162; 188pp; English.

XX

CC The present invention describes an isolated peptide (I) comprising a
CC sequence of at least four amino acids, where the peptide is a substrate
CC for conducting aspartyl protease assays. (I) has neuroprotective and
CC nootropic activities, and can be used as an inhibitor of beta-secretase
CC activity. A beta-secretase modulator from the present invention can be
CC used for inhibiting beta-secretase activity in vivo, and in the
CC manufacture of a medicament for the treatment of Alzheimer's disease.
CC Pharmaceutical compositions from the present invention can be used for
CC treating a disease or condition characterised by an abnormal beta-
CC secretase activity. (I) is useful for identifying agents that modulate
CC the activity of human Asp2 aspartyl protease (Hu-Asp2). (I) is useful as
CC a core structure to construct derivatives. ABL49914 to ABL49925 and
CC ABB06409 to ABB06593 represent sequences used in the exemplification of
CC the present invention

XX

SQ Sequence 518 AA;

Query Match 100.0%; Score 2687; DB 5; Length 518;
Best Local Similarity 100.0%; Pred. No. 8.6e-231;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSQTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA	300
Db	241	GSQTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAETIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAETIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480

Qy 481 AILLVLIVLLLLPFRCQRRPRDPEVVNDESSLVRHRWK 518
 ||||||||||||||||||||||||||||||||||||
 Db 481 AILLVLIVLLLLPFRCQRRPRDPEVVNDESSLVRHRWK 518

Search completed: March 4, 2004, 15:35:42
Job time : 107.702 secs

OM protein - protein search, using sw model

Run on: March 4, 2004, 15:31:20 ; Search time 33.6149 Seconds
 (without alignments)
 795.548 Million cell updates/sec

Title: US-09-668-314C-2
 Perfect score: 2687
 Sequence: 1 MGALARALLPLLAQWLLRA.....RPRDPEVVNDESSLVRHRWK 518

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : Issued_Patents_AA:*
 1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:*
 2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep:*
 3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep:*
 4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:*
 5: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:*
 6: /cgn2_6/ptodata/2/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2687	100.0	518	3	US-08-999-723-2	Sequence 2, Appli
2	2687	100.0	518	3	US-09-434-427-2	Sequence 2, Appli
3	2687	100.0	518	4	US-09-548-372D-2	Sequence 2, Appli
4	2687	100.0	518	4	US-09-548-367D-2	Sequence 2, Appli
5	2687	100.0	518	4	US-09-551-853D-2	Sequence 2, Appli
6	2687	100.0	518	4	US-09-215-450-19	Sequence 19, Appl
7	2395	89.1	514	3	US-09-717-432-2	Sequence 2, Appli
8	2395	89.1	514	4	US-09-912-484-2	Sequence 2, Appli
9	1186.5	44.2	501	4	US-09-713-158-2	Sequence 2, Appli
10	1185	44.1	501	4	US-09-548-372D-8	Sequence 8, Appli
11	1185	44.1	501	4	US-09-548-367D-8	Sequence 8, Appli

12	1185	44.1	501	4	US-09-551-853D-8	Sequence 8, Appli
13	1184.5	44.1	501	4	US-09-724-566A-65	Sequence 65, Appl
14	1178.5	43.9	501	4	US-09-548-372D-4	Sequence 4, Appli
15	1178.5	43.9	501	4	US-09-548-367D-4	Sequence 4, Appli
16	1178.5	43.9	501	4	US-09-551-853D-4	Sequence 4, Appli
17	1178.5	43.9	501	4	US-09-724-566A-2	Sequence 2, Appli
18	1175	43.7	488	4	US-09-604-608-2	Sequence 2, Appli
19	1175	43.7	503	4	US-09-604-608-3	Sequence 3, Appli
20	1172.5	43.6	501	4	US-09-009-191-2	Sequence 2, Appli
21	1171.5	43.6	456	4	US-09-724-566A-43	Sequence 43, Appl
22	1171.5	43.6	480	4	US-09-724-566A-66	Sequence 66, Appl
23	1167.5	43.4	444	4	US-09-724-566A-67	Sequence 67, Appl
24	1160.5	43.2	439	4	US-09-724-566A-69	Sequence 69, Appl
25	1160.5	43.2	774	4	US-09-009-191-4	Sequence 4, Appli
26	1139	42.4	452	4	US-09-724-566A-59	Sequence 59, Appl
27	1139	42.4	453	4	US-09-548-372D-30	Sequence 30, Appl
28	1139	42.4	453	4	US-09-548-367D-30	Sequence 30, Appl
29	1139	42.4	453	4	US-09-551-853D-30	Sequence 30, Appl
30	1139	42.4	459	4	US-09-548-372D-32	Sequence 32, Appl
31	1139	42.4	459	4	US-09-548-367D-32	Sequence 32, Appl
32	1139	42.4	459	4	US-09-551-853D-32	Sequence 32, Appl
33	1132	42.1	407	4	US-09-724-566A-58	Sequence 58, Appl
34	1132	42.1	431	4	US-09-724-566A-74	Sequence 74, Appl
35	1128	42.0	395	4	US-09-724-566A-68	Sequence 68, Appl
36	1127	41.9	433	4	US-09-548-372D-26	Sequence 26, Appl
37	1127	41.9	433	4	US-09-548-367D-26	Sequence 26, Appl
38	1127	41.9	433	4	US-09-551-853D-26	Sequence 26, Appl
39	1127	41.9	446	4	US-09-548-372D-22	Sequence 22, Appl
40	1127	41.9	446	4	US-09-548-367D-22	Sequence 22, Appl
41	1127	41.9	446	4	US-09-551-853D-22	Sequence 22, Appl
42	1127	41.9	459	4	US-09-548-372D-24	Sequence 24, Appl
43	1127	41.9	459	4	US-09-548-367D-24	Sequence 24, Appl
44	1127	41.9	459	4	US-09-551-853D-24	Sequence 24, Appl
45	1124	41.8	420	4	US-09-724-566A-60	Sequence 60, Appl

ALIGNMENTS

RESULT 1
 US-08-999-723-2
 ; Sequence 2, Application US/08999723A
 ; Patent No. 6025180
 ; GENERAL INFORMATION:
 ; APPLICANT: Powell, David J.
 ; APPLICANT: Southan, Christopher
 ; APPLICANT: Chapman, Conrad G.
 ; APPLICANT: Evans, Joanne R.
 ; TITLE OF INVENTION: ASP1
 ; FILE REFERENCE: GH70262
 ; CURRENT APPLICATION NUMBER: US/08/999,723A
 ; CURRENT FILING DATE: 1997-10-06
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 518
 ; TYPE: PRT

; ORGANISM: Homo sapiens
US-08-999-723-2

Query Match 100.0%; Score 2687; DB 3; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.2e-243;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGFNTSFLVNIATI 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGFNTSFLVNIATI 180

Qy    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy    241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300

Qy    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360

Qy    361 YLRDENS SRSFRTILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 YLRDENS SRSFRTILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

Qy    421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPI LWIVSYALMSVCG 480
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPI LWIVSYALMSVCG 480

Qy    481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518
        ||||||||||||||||||||||||||||||||||||||||
Db    481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518
```

RESULT 2

US-09-434-427-2

; Sequence 2, Application US/09434427

; Patent No. 6162630

; GENERAL INFORMATION:

; APPLICANT: POWELL, DAVID J.

; APPLICANT: SOUTHAN, CHRISTOPHER

; APPLICANT: CHAPMAN, CONRAD G.

; APPLICANT: EVANS, JOANNE R.

; TITLE OF INVENTION: ASP1

; FILE REFERENCE: GH-70262-D1

; CURRENT APPLICATION NUMBER: US/09/434,427

; CURRENT FILING DATE: 1999-11-04

; EARLIER APPLICATION NUMBER: US 08/999,723
; EARLIER FILING DATE: 1997-10-06
; EARLIER APPLICATION NUMBER: UK 9626022.9
; EARLIER FILING DATE: 1996-12-14
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-434-427-2

Query Match 100.0%; Score 2687; DB 3; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.2e-243;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAANFLAMVDNLQDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAANFLAMVDNLQDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDFGTGSQLACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDFGTGSQLACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLVLLIVLLLLPFRQCRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLVLLIVLLLLPFRQCRRPRDPEVVNDESSLVRHRWK	518

RESULT 3
US-09-548-372D-2
; Sequence 2, Application US/09548372D
; Patent No. 6420534

```

; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR
AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-2

```

```

Query Match          100.0%;  Score 2687;  DB 4;  Length 518;
Best Local Similarity 100.0%;  Pred. No. 2.2e-243;
Matches 518;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
        |||
Db      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
        |||
Db     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
        |||
Db    121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180

Qy    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLVTQANIPNVFSMQMCGAGLPVA 240
        |||
Db    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLVTQANIPNVFSMQMCGAGLPVA 240

Qy    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA 300
        |||
Db    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA 300

Qy    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360
        |||
Db    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360

Qy    361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
        |||
Db    361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

Qy    421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

```

```

                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy      481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518
                ||||||||||||||||||||||||||||||||||||||||
Db      481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518

```

RESULT 4

US-09-548-367D-2

; Sequence 2, Application US/09548367D

; Patent No. 6440698

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR
AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280H

; CURRENT APPLICATION NUMBER: US/09/548,367D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 518

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-548-367D-2

Query Match 100.0%; Score 2687; DB 4; Length 518;

Best Local Similarity 100.0%; Pred. No. 2.2e-243;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180

Qy    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy    241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300

```

```

          |||||||
Db      241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300

Qy      301 IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
          |||||||
Db      301 IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360

Qy      361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIED 420
          |||||||
Db      361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIED 420

Qy      421 RAQKRVGFAASPCAEIAGA AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
          |||||||
Db      421 RAQKRVGFAASPCAEIAGA AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy      481 AILLVLIVLLLLLPFRCQRRPRDPEVVNDESSLVRHRWK 518
          |||||||
Db      481 AILLVLIVLLLLLPFRCQRRPRDPEVVNDESSLVRHRWK 518

```

RESULT 5

US-09-551-853D-2

; Sequence 2, Application US/09551853D

; Patent No. 6500667

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280L

; CURRENT APPLICATION NUMBER: US/09/551,853D

; CURRENT FILING DATE: 2000-04-18

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 518

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-551-853D-2

Query Match 100.0%; Score 2687; DB 4; Length 518;

Best Local Similarity 100.0%; Pred. No. 2.2e-243;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      1 MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP GTPAERHADGL 60
          |||||||
Db      1 MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP GTPAERHADGL 60

```

```

Qy      61 ALALEPALASPAGAAANFLAMVDNLQGD SGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

```

```

      |||
Db      61 ALALEPALASPAGAAFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy      121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI 180
      |||
Db      121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI 180

Qy      181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
      |||
Db      181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy      241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
      |||
Db      241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300

Qy      301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
      |||
Db      301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360

Qy      361 YLRDENSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVM EGFYVIFD 420
      |||
Db      361 YLRDENSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVM EGFYVIFD 420

Qy      421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPI LWIVSYALMSVCG 480
      |||
Db      421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPI LWIVSYALMSVCG 480

Qy      481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518
      |||
Db      481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518

```

RESULT 6

US-09-215-450-19

; Sequence 19, Application US/09215450

; Patent No. 6635748

; GENERAL INFORMATION:

; APPLICANT: Giese, Klaus

; APPLICANT: Xin, Hong

; TITLE OF INVENTION: METASTATIC BREAST AND COLON CANCER REGULATED GENES

; FILE REFERENCE: 1451.100 / 210030.447

; CURRENT APPLICATION NUMBER: US/09/215,450

; CURRENT FILING DATE: 1998-12-17

; NUMBER OF SEQ ID NOS: 27

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 19

; LENGTH: 518

; TYPE: PRT

; ORGANISM: human

US-09-215-450-19

```

Query Match          100.0%;  Score 2687;  DB 4;  Length 518;
Best Local Similarity 100.0%;  Pred. No. 2.2e-243;
Matches 518;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPT PGP GTPAERHADGL 60
      |||

```

Db 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGPGPATPAERHADGL 60
 Qy 61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
 ||||||||||||||||||
 Db 61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
 Qy 121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTI PKGFNTSFLVNIATI 180
 ||||||||||||||||||
 Db 121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTI PKGFNTSFLVNIATI 180
 Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 ||||||||||||||||||
 Db 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 Qy 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 ||||||||||||||||||
 Db 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 Qy 301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 ||||||||||||||||||
 Db 301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 Qy 361 YLRDENSSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 ||||||||||||||||||
 Db 361 YLRDENSSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 Qy 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 ||||||||||||||||||
 Db 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 Qy 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518
 ||||||||||||||||||
 Db 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518

RESULT 7

US-09-717-432-2

; Sequence 2, Application US/09717432

; Patent No. 6291223

; GENERAL INFORMATION:

; APPLICANT: ZHU, YUAN

; APPLICANT: LI, XIAOTONG

; APPLICANT: CHRISTIE, GARY

; APPLICANT: POWELL, DAVID J.

; TITLE OF INVENTION: Mouse Aspartic Secretase-1 (mASPl)

; FILE REFERENCE: GP-70663

; CURRENT APPLICATION NUMBER: US/09/717,432

; CURRENT FILING DATE: 2000-11-21

; PRIOR APPLICATION NUMBER: 60/166,974

; PRIOR FILING DATE: 1999-11-23

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 2

; LENGTH: 514

; TYPE: PRT

; ORGANISM: MUS MUSCULUS

US-09-717-432-2

Query Match 89.1%; Score 2395; DB 3; Length 514;
 Best Local Similarity 88.6%; Pred. No. 5.6e-216;
 Matches 459; Conservative 20; Mismatches 35; Indels 4; Gaps 1;

```

Qy      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
      |||| |||| :||||| | | ||||| ||||:| | || | : || || | ||||
Db      1 MGALLRALLLLVLQAQWLLSAVPALAPAPFTLPLQVAGATNHRASAVPGLGTPELPRADGL 60

Qy     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
      ||||| | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 ALALEPVRAT----ANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKVQILVDTGSSNFAVAG 116

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGFNTSFLVNIATI 180
      ||||| ||||:| |||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    117 APHSYIDTYFDSESSSTYHSGKFDVTVKYTQGSWTGFVGEDLV TIPKGFNSSFLVNIATI 176

Qy    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    177 FESENFFLPGIKWNGILGLAYAALAKPSSSLETFFDSLVAQAKIPDIFSMQMCGAGLPVA 236

Qy    241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    237 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQNLNLD CREYNADKA 296

Qy    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    297 IVDSGTTLLRLPQKVFDVAVVEAVARTSLIPEFSDGFWTGAQLACWTNSETPWAYFPKISI 356

Qy    361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
      ||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    357 YLRDENASRSFRITILPQLYIQPMMGAGFNIECYRFGISSSTNALVIGATVMEGFYVIFD 416

Qy    421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
      |||:|||| ||||| | ||||| ||||| :||| ||||| :||| ||||| |||||
Db    417 RAQRRVGFAVSPCAIEIEGTTVSEISGPFSTEDIASNCVPAQALNEPILWIVSYALMSVCG 476

Qy    481 AILLVLIVLLLLLPPRCQRRPRDPEVVNDESSLVRHRWK 518
      ||||| :||| | : ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    477 AILLVLIVLLLLLPLHCRHAPRDPEVVNDESSLVRHRWK 514

```

RESULT 8

US-09-912-484-2

; Sequence 2, Application US/09912484

; Patent No. 6358725

; GENERAL INFORMATION:

; APPLICANT: Christie, Gary

; APPLICANT: Li, Xiaotong

; APPLICANT: Powell, David J.

; APPLICANT: Zhu, Yuan

; TITLE OF INVENTION: Mouse Aspartic Secretase-1 (mASP1)

; FILE REFERENCE: GP-70663-D1

; CURRENT APPLICATION NUMBER: US/09/912,484

; CURRENT FILING DATE: 2001-07-25

; PRIOR APPLICATION NUMBER: 60/166,974

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: 09/717,432
; PRIOR FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 514
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-09-912-484-2

Query Match 89.1%; Score 2395; DB 4; Length 514;
Best Local Similarity 88.6%; Pred. No. 5.6e-216;
Matches 459; Conservative 20; Mismatches 35; Indels 4; Gaps 1;

```
Qy      1 MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
      |||| ||||| :||||| | | ||||| |||||:|| ||| : || ||| ||||
Db      1 MGALLRALLLLVLAQWLLSAVPALAPAPFTLPLQVAGATNHRASAVPGLGTPELPRADGL 60

Qy     61 ALALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
      ||||| | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db     61 ALALEPVRAT----ANFLAMVDNLQGDSGRGYYLEMLIGTPPQKVQILVDTGSSNFAVAG 116

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI 180
      ||||| |||||:| |||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    117 APHSYIDTYFDSESSSTYHSKGFDTVVKYTQGSWTGFVGEDLV TIPKGFNSSFLVNIATI 176

Qy    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    177 FESENFFLPGIKWNGILGLAYAALAKPSSSLETFFDSLVAQAKIPDIFSMQMCGAGLPVA 236

Qy    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    237 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQNLNLD CREYNADKA 296

Qy    301 IVDSGTTLLRLPQKVFD AVEAVARASLIPEFS DGFWTGSQLACWTNSETPWSYFPKISI 360
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    297 IVDSGTTLLRLPQKVFD AVEAVARTSLIPEFS DGFWTGAQLACWTNSETPWAYFPKISI 356

Qy    361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
      |||||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    357 YLRDENASRSFRITILPQLYIQPMMGAGFN YECYRFGISSSTNALVIGATVMEGFYVIFD 416

Qy    421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
      |||:||||| ||||| | ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    417 RAQRRVGFAVSPCAIEIEGTTVSEISGPFSTEDIASNCVPAQALNEPILWIVSYALMSVCG 476

Qy    481 AILLVLIVLLLLLPFRCQRRPRDPEVVNDESSLVRHRWK 518
      |||||:||||| | : ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db    477 AILLVLILLLLLPLHCRHAPRDPEVVNDESSLVRHRWK 514
```

RESULT 9

US-09-713-158-2

; Sequence 2, Application US/09713158
; Patent No. 6361975
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN

```
; APPLICANT: LI, XIAOTONG
; APPLICANT: POWELL, DAVID J.
; APPLICANT: CHRISTIE, GARY
; TITLE OF INVENTION: MOUSE ASPARTIC SECRETASE-2 (MASP-2)
; FILE REFERENCE: GP-70660
; CURRENT APPLICATION NUMBER: US/09/713,158
; CURRENT FILING DATE: 2000-11-15
; PRIOR APPLICATION NUMBER: 60/165,800
; PRIOR FILING DATE: 1999-11-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 501
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-09-713-158-2
```

```
Query Match          44.2%; Score 1186.5; DB 4; Length 501;
Best Local Similarity 46.0%; Pred. No. 1.4e-102;
Matches 238; Conservative 83; Mismatches 167; Indels 29; Gaps 8;
```

```
Qy      7 ALLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTTPGPCTPAERHADGLA 61
      | ||| | :      : || |      ||| |      | | | | :
Db      2 AQALPWLLLVW---GSGMLPAQGTHLGIRLPLRSLA-----GPPLGLRLPRETDEES-- 51

Qy     62 LALEPALASPAGAAFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGT 121
      |      :| : ||||| :| ||: ||| :|| :| :|| | ||||| |||||
Db     52 -----EEPGRGGSFVEMVDNLRGKSGQGYVEMTVGSPPQTLNILVDTGSSNFAVGAA 104

Qy    122 PHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIF 181
      || : : | : : |||||      | | |||| | | : | ||| :|| | | : ||| |
Db    105 PHPFLHRYYQRQLSSTYRDLRKGVPYPTQGWEGELGTDLVSIHPGPNVTVRANIAAIT 164

Qy    182 ESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPV-- 239
      || : || : | | ||||| || :| : ||| ||||| | : ||| :|| :| :||| | :
Db    165 ESDKFFINGSNWEGLGLAYAEIARPDDSLEPFFDSLVKQTHIPNIFSLQLCGAGFPLNQ 224

Qy    240 -AGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNAD 298
      : ||| : : ||| : ||| | : ||||| : ||||| : | : : || | | : || : || |
Db    225 TEALASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYD 284

Qy    299 KAIVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKI 358
      | : ||||| ||||| : ||| : | : : : || | | || | | ||| : || |
Db    285 KSIVDSGTTLRLRPKKVFEEAAVKSIIKAAASSTEKFPDGFVLGEQLVCWQAGTTPWNI FPVI 344

Qy    359 SIYLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGFYV 417
      | : || | | : : ||||| ||||| | : ||| : | : | : | : |||||
Db    345 SLYLMGEVTNQSFRTILPQQYL RPVEDVATSQDDCYKFAVSQSSTGTVMGAVIMEGFYV 404

Qy    418 IFDRAQKRVGFAASPCAIEIAGA AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMS 477
      : ||||| : ||| | |      : : ||| | | :      | | : : | : :
Db    405 VFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAA 464

Qy    478 VCGAILLVLIIVLLLLPFRQCR--RPRDPEVVNDESSL 512
      : | | : : : | : : : || | | | : : : | | |
Db    465 IC--ALFMLPLCLMVCQWRCLRLRHQHDDFADDISLL 500
```

RESULT 10
 US-09-548-372D-8
 ; Sequence 8, Application US/09548372D
 ; Patent No. 6420534
 ; GENERAL INFORMATION:
 ; APPLICANT: GURNEY ET AL.
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR
 AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: 29915/6280I
 ; CURRENT APPLICATION NUMBER: US/09/548,372D
 ; CURRENT FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: US 60/155,493
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: US 09/404,133
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: US 60/101,594
 ; PRIOR FILING DATE: 1998-09-24
 ; NUMBER OF SEQ ID NOS: 73
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-548-372D-8

Query Match 44.1%; Score 1185; DB 4; Length 501;
 Best Local Similarity 46.0%; Pred. No. 1.9e-102;
 Matches 237; Conservative 83; Mismatches 169; Indels 26; Gaps 7;

Qy	9	LLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLALA	63
		: : :	
Db	1	MAPALHWLLLWVGSGMLPAQGTHLGIRLPLRSGLA-----GPPLGLRLPRETDEES----	51
Qy	64	LEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPH	123
		: : : : : :	
Db	52	-----EEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFAVGAAPH	106
Qy	124	SYIDTYFDTERSSTYRSKGFVTVKYTQGSWTGFVGEDLVTPKGFNTSFLVNIATIFES	183
		: : : : : :	
Db	107	PFLHRYYQRQLSSTYRDLRKG VYPYTQGWEGELGTDLVSI PHGPNVTVRANIAAITES	166
Qy	184	ENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPV---A	240
		: : : : : : :	
Db	167	DKFFINGSNWEGLGLAYAEIARPDDSLEPFFDSLVKQTHIPNIFSLQLCGAGFPLNQTE	226
Qy	241	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
		: : : : : : : : : : :	
Db	227	ALASVGGSMIIGGIDHSLYTGSLWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNYDKS	286
Qy	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDFWTGSQLACWTNSETPWSYFPKISI	360
		: : : : :	
Db	287	IVDSGTTNLRPLPKVFEAAVKSIIKAASSTEKFPDGFVLGEQLVCWQAGTTPWNIFFVISL	346

Qy 361 YLRDENSRSFRITILPQLYIQPMGAGLNY-ECYRFGISPSTNALVIGATVMEGFYVIF 419
 || | ::||| ||| |::| : : ||:| :| | : ||| :||| ||:|
 Db 347 YLMGEVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIMEGFYVVF 406

Qy 420 DRAQKRVGFAASPCAETIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVC 479
 |||:|:| | | : : ||| | | : | | : :| : :|
 Db 407 DRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAAIC 466

Qy 480 GAILLVLLVLLLLPFCRQR--RPRDPEVVNDESSL 512
 | : : : | : : || | | : : : | | |
 Db 467 -ALFMLPLCLMVCQWRCLRLRHQHDDEADDISLL 500

RESULT 11

US-09-548-367D-8

; Sequence 8, Application US/09548367D

; Patent No. 6440698

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280H

; CURRENT APPLICATION NUMBER: US/09/548,367D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 501

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-548-367D-8

Query Match 44.1%; Score 1185; DB 4; Length 501;

Best Local Similarity 46.0%; Pred. No. 1.9e-102;

Matches 237; Conservative 83; Mismatches 169; Indels 26; Gaps 7;

Qy 9 LLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLALA 63
 : | | | | : | | | | | | | | :
 Db 1 MAPALHWLLWLWVGSGMLPAQGTHLGIRLPLRSGLA-----GPPLGLRLPRETDEES----- 51

Qy 64 LEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPH 123
 | :|: |||||:| ||:||||:| :|:| | | ||||| || ||
 Db 52 -----EEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPQTLNILVDTGSSNFAVGAAPH 106

Qy 124 SYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGFNTSFLVNIATIFES 183
 : : | : : |||| | | |||| | | : | |||:| | | : ||| | ||
 Db 107 PFLHRYYYQRQLSSTYRDLRKGVYVPYTQGWEGELGTDLV SIPHGPVTVRANIAAITES 166

Qy 184 ENFFLPGIKWNGILGLAYATLAKPSSSLETFFDLVTQANIPNVFSMQMCGAGLPV---A 240
: ||: | | ||||| :|:| ||| ||||| | :||:|:|:| ||| |:

Db 167 DKFFINGSNWEGLGLAYAEIARPDDSLEPFDSLQTHIPNIFSLQLCGAGFPLNQTE 226

Qy 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
: |||:::| ||| | :|||: |||:| :::|| | | :||:| ||| |:

Db 227 ALASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKS 286

Qy 301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360
||||| |||:|:| | ::: || :| ||| | || | |||: || |:

Db 287 IVDSGTTNLRPLPKVFEEAAVKSIIKAASSTEKFPDGEWLGEQLVCWQAGTTPWNIFPVISL 346

Qy 361 YLRDENSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGFYVIF 419
|| | :::| ||||| | :|: : :||:| :| |: | :|| :|||:|

Db 347 YLMGEVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIMEGFYVVF 406

Qy 420 DRAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVC 479
|||:|:| || | : : ||| | |: | | :| : :|

Db 407 DRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAAIC 466

Qy 480 GAILLVLLVLLLLPFCRQR--RPRDPEVVNDESSL 512
|: : : |:: :|| | | : : :| | |

Db 467 -ALFMLPLCLMVCQWRCLRLRHQHDDFADDISLL 500

RESULT 12

US-09-551-853D-8

; Sequence 8, Application US/09551853D

; Patent No. 6500667

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280L

; CURRENT APPLICATION NUMBER: US/09/551,853D

; CURRENT FILING DATE: 2000-04-18

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8

; LENGTH: 501

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-551-853D-8

Query Match 44.1%; Score 1185; DB 4; Length 501;

Best Local Similarity 46.0%; Pred. No. 1.9e-102;

Matches 237; Conservative 83; Mismatches 169; Indels 26; Gaps 7;

Qy 9 LLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLALA 63
: | | | | : | | | | | | | | :
Db 1 MAPALHWLLWVGSGMLPAQGTHLGIRLPLRSGLA-----GPPLGLRLPRETDEES----- 51

Qy 64 LEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPH 123
| : | | | | : | | | | : | | | | | | | | | |
Db 52 -----EEPGRRGSEFVEMVDNLRGKSGQGYVEMTVGSPPQTLNILVDTGSSNFAVGAAPH 106

Qy 124 SYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVITPKGENTSFLVNIATIFES 183
: : | : : | | | | | | | | | | : | | | | : | | | | |
Db 107 PFLHRYYYQRQLSSTYRDLRKGVYVPYTQGWEGELGTDLVSIHPGPNVTVRANIAAITES 166

Qy 184 ENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPV---A 240
: | | : | | | | | | : | | | | | | : | | | : | | | | :
Db 167 DKFFINGSNWEGLGLAYAEIARPDSDLPEFFDSLQTHIPNIFSLQCGAGFPLNQTE 226

Qy 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA 300
: | | : : | | : | | | : | | : : | : : | | | : | | : | |
Db 227 ALASVGGSMIIGGIDHSLYTGSLWYTPIRREWYEVIIVRVEINGQDLKMDCKEYNYDKS 286

Qy 301 IVDSGTTLLRLPQKVFDAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360
| | | | | | | | : | | : | | : | | : | | | | | | : | | : | |
Db 287 IVDSGTTNLRPLPKVFEEAAVKSIIKAASSTEKFPDGFVLGEQLVCWQAGTTPWNIFPVISL 346

Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYREGISPSTNALVIGATVMEGFYVIF 419
| | | : : | | | | | | | : : | : : | : | : | : | : | : | : |
Db 347 YLMGEVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIMEGFYVVF 406

Qy 420 DRAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPIILWIVSYALMSVC 479
| | : | : | | | | : : | | | | : | | : | : | : : | : : |
Db 407 DRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAAIC 466

Qy 480 GAILLVLLVLLLLPFRQCR--RPRDPEVVNDESSL 512
| : : : | : : | | | : : : | | |
Db 467 -ALFMLPLCLMVCQWRCLRLRHQHDDFADDISLL 500

RESULT 13

US-09-724-566A-65

; Sequence 65, Application US/09724566A

; Patent No. 6627739

; GENERAL INFORMATION:

; APPLICANT: Anderson, John P.

; APPLICANT: Basi, Gurigbal

; APPLICANT: Doane, Minh Tam

; APPLICANT: Frigon, No. 6627739mand

; APPLICANT: John, Varghese

; APPLICANT: Power, Michael

; APPLICANT: Sinha, Sukanto

; APPLICANT: Tatsuno, Gwen

; APPLICANT: Tung, Jay

; APPLICANT: Wang, Shuwen

; APPLICANT: McConlogue, Lisa

; TITLE OF INVENTION: Beta-Secretase Enzyme Compositions and

; TITLE OF INVENTION: Methods

; FILE REFERENCE: 228-US-NEWC2

; CURRENT APPLICATION NUMBER: US/09/724,566A

```
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: US 09/501,708
; PRIOR FILING DATE: 2000-02-10
; PRIOR APPLICATION NUMBER: 60/119,571
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/139,172
; PRIOR FILING DATE: 1999-06-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-724-566A-65
```

```
Query Match          44.1%; Score 1184.5; DB 4; Length 501;
Best Local Similarity 45.9%; Pred. No. 2.1e-102;
Matches 237; Conservative 84; Mismatches 170; Indels 25; Gaps 7;
```

```
Qy      9 LLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLALA 63
      : | |   | |   : | |   |   | | |   |   |   | |   :
Db      1 MAPALHWLLLWVGSGMLPAQGTHLGIRLPLRSGLA-----GPPLGLRLPRETDEES---- 51

Qy     64 LEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPH 123
      |   : | : | | | | : | | : | | : | | : | | | | | | | | | | | |
Db     52 -----EEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPQTLNILVDTGSSNFAVGAAPH 106

Qy    124 SYIDTYFDTERSSTYRSKGFDTVKYTGQSWTGFVGEDLVITPKGFNTSFLVNIATIFES 183
      : : | : | : | | | | | | | | | | | | | | | | | | | | | |
Db    107 PFLHRYYYQRQLSSTYRDLRKGVYVPYTQGWEGELGTDLVSI PHGPNVTVRANIAAITES 166

Qy    184 ENFFLPGIKWNIGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPV---A 240
      : | | : | | | | | | | | | | | | | | | | | | | | | | | |
Db    167 DKFFINGSNWEIGILGLAYAEIARPDSDLPEFFDSLQTHIPNIFSLQLCGAGFPLNQTE 226

Qy    241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
      : | | | : : | | | : | | | | : | | | | : | | | : | | | | | :
Db    227 ALASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKS 286

Qy    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360
      | | | | | | | | | | : | | | : | | | | | | | | | | | | | | |
Db    287 IVDSGTTNLRPLPKVFEEAAVKSIIKAASSTEKFPDGEFWLGEQLVCWQAGTTPWNIFPVISL 346

Qy    361 YLRDENSRRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGFYVIF 419
      | | | | : : | | | | | | | | | | : | | | : | | | : | | | | | :
Db    347 YLMGEVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIMEGFYVVF 406

Qy    420 DRAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPIILWIVSYALMSVC 479
      | | | : | | : | | | | | | | : : | | | | | : | | | : : |
Db    407 DRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAAIC 466

Qy    480 GAILLVLLVLLLLPFCRQRRPR-DPEVVNDESSLVR 514
      | : : : | : : : | | | | | : | : | | : :
Db    467 -ALFMLPLCLMVCQWRCLRLRHQHDDFGDDISLLK 501
```

US-09-548-372D-4
 ; Sequence 4, Application US/09548372D
 ; Patent No. 6420534
 ; GENERAL INFORMATION:
 ; APPLICANT: GURNEY ET AL.
 ; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR
 AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: 29915/6280I
 ; CURRENT APPLICATION NUMBER: US/09/548,372D
 ; CURRENT FILING DATE: 2000-04-12
 ; PRIOR APPLICATION NUMBER: US 60/155,493
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: US 09/404,133
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: PCT/US99/20881
 ; PRIOR FILING DATE: 1999-09-23
 ; PRIOR APPLICATION NUMBER: US 60/101,594
 ; PRIOR FILING DATE: 1998-09-24
 ; NUMBER OF SEQ ID NOS: 73
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 4
 ; LENGTH: 501
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-548-372D-4

Query Match 43.9%; Score 1178.5; DB 4; Length 501;
 Best Local Similarity 46.2%; Pred. No. 7.8e-102;
 Matches 240; Conservative 82; Mismatches 164; Indels 33; Gaps 9;

Qy	7	ALLLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLA	61
		: :	
Db	2	AQALPWLLLWM---GAGVLPAGHTQHGI RLPLRSGLG-----GAPL-----GLR	42
Qy	62	LALE--PALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVA	119
		: : :	
Db	43	LPRETDEEPEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPQTLN ILVDTGSSNFAVG	102
Qy	120	GTPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIAT	179
		: :	
Db	103	AAPHPFLHRY YQRLSSTYRDLRKG VYPYTQGWEGELGTDLVSI PHGPNVTVRANIAA	162
Qy	180	IFESENFFLPGLIKWNGILGLAYATLAKPSSSLETF FDSLVTQANIPNVF'SMQMCGAGLPV	239
		: : : :	
Db	163	ITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDSL VKQTHVPNLFSLQLCGAGFPL	222
Qy	240	AGS---GTNGGSLVLGGIEPSLYKGD I WYTPIKEEWY YQIEILKLEIGGQSLNLD CREYN	296
		: : : : :	
Db	223	NQSEVLASVGGS MIIGGIDHSLYTGSLWYTPIRREWY YEVII VRVEINGQDLKMDCKEYN	282
Qy	297	ADKAIVDSGTTLLRLPQKVFD AVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFP	356
		: : :	
Db	283	YDKSIVDSGTTNLR LPKKVFEAAVKS IKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFP	342
Qy	357	KIS IYLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGF	415
		: : : : : :	

```

Db      343 VISLYLMGEVTNQSFRTITLPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGF 402
Qy      416 YVIFDRAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYAL 475
        ||:||||:||:|||| | | : : ||| | |: | | ::| :
Db      403 YVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDESTLMTIAYVM 462
Qy      476 MSVCGAILLVLLIVLLLLPFCQR--RPRDPEVVNDESSL 512
        ::| |: :: : |: : :|| | | : : :| | |
Db      463 AAIC-ALEMLPLCLMVCQWRCLRCLRQQHDDFADDISLL 500

```

RESULT 15

US-09-548-367D-4

```

; Sequence 4, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR
AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-4

```

```

Query Match          43.9%; Score 1178.5; DB 4; Length 501;
Best Local Similarity 46.2%; Pred. No. 7.8e-102;
Matches 240; Conservative 82; Mismatches 164; Indels 33; Gaps 9;

```

```

Qy      7 ALLLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTTPGPGTPAERHADGLA 61
        | || | |: : || | ||| | | ||
Db      2 AQALPWLLLWM---GAGVLPAGHTQHGIRLPLRSLG-----GAPL-----GLR 42
Qy      62 LALE--PALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVA 119
        | | : |: ||||:| ||:||||:| |:||| | |||||
Db      43 LPRETDEEPEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFAVG 102
Qy      120 GTPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTPKGFNTSFLVNIAT 179
        || :: |: : |||| | | |||| | | :| |||:| | | : |||
Db      103 AAPHPFLHRYYYQRQLSSTYRDLRKGVYVPYTQGWEGELGTDLVSIHPGPNVTVRANIAA 162
Qy      180 IFESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDLSLVTQANIPNVFSMQMCGAGLPV 239
        | ||: ||: | | ||||| ||:| ||| ||||| | ::||:| |:| ||| | :

```

Db 163 ITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFDSLQTHVPNLFSLQLCGAGFPL 222
 Qy 240 AGS---GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYN 296
 | : |||::|||: ||| | :||| | :|||: |::|| || | :||:|
 Db 223 NQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYN 282
 Qy 297 ADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFP 356
 ||:||||| | |||:| | :| :| | | | | | | | :| |
 Db 283 YDKSIVDSGTTNLR LRPKKVFEEAAVKS IKAASSTEKFDPDGEWLGEQLVCWQAGTTPWNIFP 342
 Qy 357 KISIYLRDENS SRSFRTILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGF 415
 ||:| | | :::| | | | | | | :| : : :| :| | | : | :| :| | |
 Db 343 VISLYLMGEVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGF 402
 Qy 416 YVIEDRAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYAL 475
 ||:| | | :| :| | | | : : | | | | : | | :| :
 Db 403 YVVEDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDESTLMTIAYVM 462
 Qy 476 MSVCGAILLV LIVLLLLPFCQR--RPRDPEVVNDESSL 512
 ::| | : : : | : : | | | | : : : | | |
 Db 463 AAIC-ALFMLPLCLMVCQWRCLRCLRQQHDDFADDISLL 500

Search completed: March 4, 2004, 15:42:13
 Job time : 34.6149 secs

OM protein - protein search, using sw model

Run on: March 4, 2004, 15:39:01 ; Search time 57.8617 Seconds
(without alignments)
1890.324 Million cell updates/sec

Title: US-09-668-314C-2
Perfect score: 2687
Sequence: 1 MGALARALLPLLAQWLLRA.....RPRDPEVVNDESSLVRHRWK 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

		%					
Result		Query					
No.	Score	Match	Length	DB	ID	Description	

OM protein - protein search, using sw model

Run on: March 4, 2004, 15:30:05 ; Search time 28.1043 Seconds
 (without alignments)
 1772.942 Million cell updates/sec

Title: US-09-668-314C-2
 Perfect score: 2687
 Sequence: 1 MGALARALLPLLAQWLLRA.....RPRDPEVVNDESSLVRHRWK 518

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries

Database : PIR_78:*
 1: pir1:*
 2: pir2:*
 3: pir3:*
 4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1178.5	43.9	501	2 A59090	aspartic proteinas
2	367.5	13.7	383	2 JC7573	pepsinogen C - Afr
3	363.5	13.5	377	1 PEMQCJ	gastricsin (EC 3.4
4	355.5	13.2	384	2 A39314	gastricsin (EC 3.4
5	355	13.2	389	2 JE0371	pepsin C (EC 3.4.2
6	353	13.1	388	2 A29937	gastricsin (EC 3.4
7	351.5	13.1	388	2 JC7246	pepsinogen C - com
8	324.5	12.1	394	2 B43356	gastricsin (EC 3.4
9	320	11.9	385	2 JC7575	pepsinogen A - bul
10	320	11.9	402	1 REMSK	renin (EC 3.4.23.1
11	313.5	11.7	509	2 S66516	oryzasin (EC 3.4.2
12	313	11.6	392	1 A24608	gastricsin (EC 3.4
13	310	11.5	383	2 A41443	pepsin (EC 3.4.23.

14	308.5	11.5	412	1	KHHUD	cathepsin D (EC 3.
15	306.5	11.4	410	1	KHMSD	cathepsin D (EC 3.
16	305.5	11.4	401	1	REMSS	renin (EC 3.4.23.1
17	305	11.4	384	2	JC7574	pepsinogen A - Afr
18	305	11.4	407	1	KHRTD	cathepsin D (EC 3.
19	302	11.2	405	2	A25379	saccharopepsin (EC
20	301.5	11.2	398	2	S66465	cathepsin E (EC 3.
21	300.5	11.2	387	2	C38302	pepsin (EC 3.4.23.
22	299	11.1	398	2	I51185	cathepsin D (EC 3.
23	298.5	11.1	387	2	D38302	pepsin (EC 3.4.23.
24	298.5	11.1	400	2	I47099	renin (EC 3.4.23.1
25	297	11.1	388	1	PEHU	pepsin A (EC 3.4.2
26	296	11.0	388	2	A30142	pepsin A (EC 3.4.2
27	296	11.0	388	2	B30142	pepsin A (EC 3.4.2
28	294.5	11.0	388	1	S19684	pepsin A (EC 3.4.2
29	292	10.9	506	2	T07915	probable aspartic
30	291	10.8	388	1	S19682	pepsin A (EC 3.4.2
31	291	10.8	402	1	RERTK	renin (EC 3.4.23.1
32	291	10.8	406	1	REHUK	renin (EC 3.4.23.1
33	290.5	10.8	396	2	S36865	cathepsin E (EC 3.
34	289	10.8	387	2	E38302	pepsin (EC 3.4.23.
35	288	10.7	387	2	B38302	pepsin (EC 3.4.23.
36	288	10.7	388	1	PEMQAJ	pepsin A (EC 3.4.2
37	287.5	10.7	632	2	T45858	hypothetical prote
38	287	10.7	391	2	A43356	cathepsin E (EC 3.
39	287	10.7	396	2	A34401	cathepsin E (EC 3.
40	286.5	10.7	334	2	JC4870	pepsin A (EC 3.4.2
41	286	10.6	382	1	PECH	pepsin A (EC 3.4.2
42	286	10.6	388	1	PEMQAR	pepsin A (EC 3.4.2
43	285.5	10.6	387	2	JC7245	pepsinogen A - com
44	285	10.6	396	2	T47207	aspartic proteinas
45	284.5	10.6	386	1	PEPG	pepsin A (EC 3.4.2

ALIGNMENTS

RESULT 1

A59090

aspartic proteinase (EC 3.4.23.-) BACE precursor - human

N;Alternate names: beta-secretase; beta-site APP cleaving enzyme

C;Species: Homo sapiens (man)

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-May-2000

C;Accession: A59090

R;Vassar, R.; Bennett, B.D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E.A.; Denis, P.;

Teplov, D.B.; Ross, S.; Amarante, P.; Loeloff, R.; Luo, Y.; Fisher, S.; Fuller,

J.; Edenson, S.; Lile, J.; Jarosinski, M.A.; Biere, A.L.; Curran, E.; Burgess,

T.; Louis, J.C.; Collins, F.; Treanor, J.; Rogers, G.; Citron, M.

Science 286, 735-741, 1999

A;Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.

A;Reference number: A59090; MUID:20002972; PMID:10531052

A;Note: submitted to GenBank, September 1999

A;Accession: A59090

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-501 <VAS>

A;Cross-references: GB:AF190725; NID:g6118538; PIDN:AAF04142.1; PID:g6118539
 C;Genetics:
 A;Gene: BACE
 C;Superfamily: beta-secretase
 C;Keywords: Alzheimer's disease; aspartic proteinase; brain; glycoprotein;
 hydrolase; protein digestion; transmembrane protein; zymogen
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-45/Domain: propeptide #status predicted <PRO>
 F;46-501/Product: acid proteinase BACE #status predicted <MAT>
 F;461-477/Domain: transmembrane #status predicted <TRN>
 F;93,289/Active site: Asp #status predicted
 F;153,172,223,354/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;330-380/Disulfide bonds: #status predicted

Query Match 43.9%; Score 1178.5; DB 2; Length 501;
 Best Local Similarity 46.2%; Pred. No. 3.8e-80;
 Matches 240; Conservative 82; Mismatches 164; Indels 33; Gaps 9;

Qy	7	ALLLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLA	61
		: :	
Db	2	AQALPWLLLWM---GAGVLPAGHTQHGI RLPLRSLG-----GAPL-----GLR	42
Qy	62	LALE--PALASPAGAAFLAMVDNLQGDSSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVA	119
		: : : :	
Db	43	LPRETDEEPEEPGRGGSFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFAVG	102
Qy	120	GTPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIAT	179
		: : : : : : :	
Db	103	AAPHPFLHRYRQQLSSTYRDLRKGVPYPTQGWEGELGTDLVSIHPGNVTVRANIAA	162
Qy	180	IFESENFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLTQANIPNVFSMQMCGAGLPV	239
		: : : : : : : :	
Db	163	ITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFDSLQTHVPNLFSLQLCGAGFPL	222
Qy	240	AGS---GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDREYN	296
		: : : : : : : : : : :	
Db	223	NQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYEVIIVRVEINGQDLKMDCKEYN	282
Qy	297	ADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWNTSETPWSYFP	356
		: : : :	
Db	283	YDKSIVDSGTTNLRPLPKVFEEAAVKSIAASSTEKFPDGFVLGEQLVCWQAGTTPWNIFP	342
Qy	357	KISIIYLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGF	415
		: : : : : : : : :	
Db	343	VISLYLMGEVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGF	402
Qy	416	YVIFDRAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYAL	475
		: : : : : : : : :	
Db	403	YVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDESTLMTIAYVM	462
Qy	476	MSVCGAILLVLLIVLLLLPFRRCQR--RPRDPEVVNDESSL	512
		: : : : : : : : : : :	
Db	463	AAIC-ALFMLPLCLMVCQWRCLRLRQQHDDFADDISLL	500

pepsinogen C - African clawed frog
 N;Alternate names: progastricsin
 C;Species: *Xenopus laevis* (African clawed frog)
 C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
 C;Accession: JC7573; PC7118
 R;Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
 J. Biochem. 129, 147-153, 2001
 A;Title: Amphibian pepsinogens: Purification and characterization of *Xenopus* pepsinogens, and molecular cloning of *Xenopus* and bullfrog pepsinogens.
 A;Reference number: JC7573; MUID:21064922; PMID:11134969
 A;Contents: Stomach
 A;Accession: JC7573
 A;Molecule type: mRNA
 A;Residues: 1-383 <IKU>
 A;Cross-references: DDBJ:AB045379
 A;Accession: PC7118
 A;Molecule type: protein
 A;Residues: 17-68 <IK2>
 C;Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like activity.
 C;Genetics:
 A;Gene: PgC
 C;Superfamily: pepsin
 C;Keywords: stomach; zymogen

Query Match 13.7%; Score 367.5; DB 2; Length 383;
 Best Local Similarity 28.9%; Pred. No. 9.1e-20;
 Matches 132; Conservative 70; Mismatches 154; Indels 101; Gaps 25;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
		: :: :: : : :::	
Db	1	MKFLILALVCLQLSEGIIR-----VPLKKFKSMREVMRENGIKAPLVDPAT---KYYNQY	52
Qy	61	ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
		: :: : :	
Db	53	ATAYEP-----LSNYMDM-----SYIGEISIGTPPQNFLVLFDGTGSSNLWVAS	95
Qy	121	TPHSYIDT-----YFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSF	173
		: : :: :	
Db	96	T---YCQSQACTNHPLFNPSQSSTYSSNQQQFSLQYGTGSLTGILGYDTVTIQ-----	145
Qy	174	LVNIATIFESENFLL----PG-----IKWNGILGLAYATLAKPSSSLETFFDSLVTQANI	224
		: : :: : : : : :	
Db	146	--NVA--ISQQEFGGLSETEPGTNFVYAQFDGILGLAYPSIAVGGAT--TVMQGM--QQNL	198
Qy	225	PN--VFESMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLE	282
		: : : :: : :: : :	
Db	199	LNQPIFGFYLSGQ-----SSQNGGEVAFGGVDQNYTGTQIYWTPVTSETYWQIGIQGFS	252
Qy	283	IGGQSLNLDCREYNADKAIVDSGTTLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ	342
		: : : : : ::::: : :	
Db	253	INGQATGW-CSQ--GCQAIVDTGTSLLTAPQSVFSSLIQSIG-----AQQDQNGQYV	301
Qy	343	ACWTNSETPWSYFPKISIIYLRDENSSRSFRITILPQLYI-QPMMGAGLNYECYRFGIS--	399
		: : : : :	
Db	302	VSCSNIQN----LPTISFTI----SGVSFPLP--PSAYVLQQSSG-----YC-TIGIMPT	345

Qy 400 --PSTNA---LVIGATVMEGFYVIFDRAQKRVGFAAS 431
 || | ::| : :| ::| :|||| :
 Db 346 YLPSQNGQPLWILGDVFLREYYSVYDLGNNQVGFATA 382

RESULT 3

PEMQCJ

gastricsin (EC 3.4.23.3) precursor - Japanese macaque (fragment)

N;Alternate names: pepsin C

C;Species: *Macaca fuscata* (Japanese macaque)

C;Date: 13-Aug-1986 #sequence_revision 19-Oct-1995 #text_change 18-Jun-1999

C;Accession: S19683; A00986; A22402; S16066

R;Kageyama, T.; Tanabe, K.; Koiwai, O.

Eur. J. Biochem. 202, 205-215, 1991

A;Title: Development-dependent expression of isozymogens of monkey pepsinogens and structural differences between them.

A;Reference number: S19681; MUID:92037645; PMID:1935977

A;Accession: S19683

A;Molecule type: mRNA

A;Residues: 1-377 <KAG>

A;Cross-references: EMBL:X59754; NID:g38072; PIDN:CAA42426.1; PID:g38073

R;Kageyama, T.; Takahashi, K.

J. Biol. Chem. 261, 4406-4419, 1986

A;Title: The complete amino acid sequence of monkey progastricsin.

A;Reference number: A00986; MUID:86168133; PMID:3514597

A;Accession: A00986

A;Molecule type: protein

A;Residues: 6-330, 'V', 332-349, 'VY', 350-377 <KA2>

R;Kageyama, T.; Takahashi, K.

J. Biochem. 97, 1235-1246, 1985

A;Title: Monkey pepsinogens and pepsins. VII. Analysis of the activation process and determination of the NH2-terminal 60-residue sequence of Japanese monkey progastricsin, and molecular evolution of pepsinogens.

A;Reference number: A22402; MUID:85289106; PMID:3928607

A;Accession: A22402

A;Molecule type: protein

A;Residues: 6-65 <KA3>

C;Comment: This enzyme has more restricted specificity than pepsin A.

C;Comment: The enzyme is activated in a two-step process that gives rise to two end products. The shorter, Ser-gastricsin, is the major product.

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach

F;1-5/Domain: signal sequence (fragment) #status predicted <SIG>

F;6-377/Product: progastricsin #status experimental <ZYM>

F;6-45/Domain: activation peptide #status experimental <APT>

F;46-377/Product: Gly-gastricsin #status experimental <MIN>

F;49-377/Product: Ser-gastricsin #status experimental <MAT>

F;31-32/Cleavage site: Phe-Leu (pepsin) #status experimental

F;45-46/Cleavage site: Phe-Gly (pepsin) #status experimental

F;48-49/Cleavage site: Leu-Ser (pepsin) #status experimental

F;80,265/Active site: Asp #status predicted

F;93-98,256-260,299-332/Disulfide bonds: #status experimental

Query Match 13.5%; Score 363.5; DB 1; Length 377;

Best Local Similarity 28.9%; Pred. No. 1.8e-19;

Matches 118; Conservative 65; Mismatches 118; Indels 107; Gaps 19;

```

Qy      56 HADGLALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSN 115
      |  |::: ||          :| :|          |: |: ||||| :| |||||
Db      44 HFGDLSVSYEP-----MAYMD-----AAYFGEISIGTPPQNFLVLFDTGSSN 85

Qy     116 FAV-----AGTPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPK 167
      |          | | | | :| | | | :| :::| | | | | | :|:
Db      86 LWVPSVYCQSQACTSHS----RFPSESTYSTNGQTFSLQYGSGLTGFFGYDTLTV-- 139

Qy     168 GFNTSFLVNIATIFESENFPLPG-----IKWNGILGLAYATLAKPSSLETFFDSLVTQA 222
      | |          | | | | :|:|:|:| | | :| :| :| :| :
Db     140 ---QSIQVPNQEFGLSEN--EPGTNEVYAQFDGIMGLAYPTLSVDGAT--TAMQGMVQEG 192

Qy     223 NIPN-VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKL 281
      : : :|: :          |::|:|:| | | :| | | :| :| | :
Db     193 ALTSPIFSIVLSDQ-----QGSSGAVVFGVDSSLYTGQIYWAPVTQELYWQIGIEEF 246

Qy     282 EIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ 341
      |||: | | :|||:|:| | :|: |::|          ||:|
Db     247 LIGGQASGW-CSE--GCQAIVDTGTSLLTVPQQYMSALLQA-----TGAQ 288

Qy     342 LACWTNSETPWSYF-----PKISYLRDENSSRSFRITILPQLYIQPMMGAGLNY 391
      | : |          | :: :          : | | | | |
Db     289 -----EDEYGQFLVNCNSIQNLPTLTFII-----NGVEFPLPPSSYI-----LNN 328

Qy     392 ECY-RFGISP-----STNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
      | |: |          | :| : :| :| : | | | :
Db     329 NGYCTVGVEPTYLSAQNSQPLWILGDVFLRSYYSVYDLSNNRVGFATA 376

```

RESULT 4

A39314

gastricsin (EC 3.4.23.3) precursor - bullfrog

C;Species: Rana catesbeiana (bullfrog)

C;Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 22-Jun-1999

C;Accession: A39314

R;Yakabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Miki, K.; Kurokawa, K.; Ito, H.; Kageyama, T.; Takahashi, K.

J. Biol. Chem. 266, 22436-22443, 1991

A;Title: Purification, characterization, and amino acid sequences of pepsinogens and pepsins from the esophageal mucosa of bullfrog (Rana catesbeiana).

A;Reference number: A39314; MUID:92042186; PMID:1939266

A;Accession: A39314

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-384 <YAK>

A;Cross-references: GB:M73750; NID:g213687; PIDN:AAA49530.1; PID:g213688

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; hydrolase; protein digestion

```

Query Match          13.2%;  Score 355.5;  DB 2;  Length 384;
Best Local Similarity 26.5%;  Pred. No. 7.2e-19;
Matches 120;  Conservative 73;  Mismatches 136;  Indels 123;  Gaps 21;

```

```

Qy      23 ELAPAPFTLPLRVAAATNRVV-----APTPGPGTPAERHADGLALALEPALASPAGAAAN 76
      :|:      :||:      : |:      || | | :| : | | | | |
Db      12 QLSEGIIVKPLKKFKSMREVMRDHGIKAPVVDPAT---KYNNFATAFEP-----LAN 61

```

```

Qy      77 FLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDT-----Y 129
      :: |          || |: ||||| :| ||||| | | :| :
Db      62 YMDM-----SYIGEISIGTPPQNFLVLFDTGSSNLWV---PSTYCQSQACTNHPQ 108

Qy     130 FDTERSSTYRSKGFDTVKYTGQSWTGFVGEDLVITPKGFNTSFLVNIATIFESENFFL- 188
      |: :||:| | :::| || || :| | | | ||| : | |
Db     109 FNPSQSSSYSSNQQQFSLQYGTGSLTGILGYDTVQIQ-----NIA--ISQQEFGLS 157

Qy     189 ---PG-----IKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPN--VFSMQMCGAGLP 238
      || :::||| ||| :::| : : | : : | | :| : |
Db     158 VTEPGTNFVYAQFDGILGLAYPSIAEGGAT--TVMQGMQI-QQNLIQPLFAFYLSG---- 210

Qy     239 VAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNAD 298
      : ||| : ||:: : | | |::||: | |::| | : ||: | :
Db     211 -QQNSQNGGEVAFGGVDQNYYSGQIYWTPVTSETYWQIGIQGFSVNGQATGW-CSQ--GC 266

Qy     299 KAIVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ-----LACWTNSET 350
      : |||:| |:| | || || :::: : | :| : :|
Db     267 QGIVDTGTSLLTAPQSVFSSLMQSI-----GAQQDQNGQYAVSCSNIQSL 311

Qy     351 PWSYFP-----KISYLRDENS---SRFRITILPQLYIQPMMGAGLNYECYREGIS 399
      | | | | : | : :|| : | | | | :|
Db     312 PTISFTISGVSFPLPPSAYVLQQNSGYCTIGIMPTYLPSQNGQPLW----- 357

Qy     400 PSTNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
      ::| : :| :| :||| :
Db     358 -----ILGDVFLRQYYSVYDLGNNQVGFAAA 383

```

RESULT 5

JE0371

pepsin C (EC 3.4.23.-) precursor - chicken

N;Alternate names: pepsinogen C

C;Species: Gallus gallus (chicken)

C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000

C;Accession: JE0371

R;Sakamoto, N.; Saiga, H.; Yasugi, S.

Biochem. Biophys. Res. Commun. 250, 420-424, 1998

A;Title: Analysis of temporal expression pattern and cis-regulatory sequences of chicken pepsinogen A and C.

A;Reference number: JE0370; MUID:98440813; PMID:9753645

A;Accession: JE0371

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-389 <SAK>

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; hydrolase

Query Match 13.2%; Score 355; DB 2; Length 389;

Best Local Similarity 28.7%; Pred. No. 7.9e-19;

Matches 114; Conservative 58; Mismatches 121; Indels 104; Gaps 16;

```

Qy      75 ANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGT-----PHSYI 126
      :|| : | : || |: ||||| :| ||||| | | | :
Db      56 SNFATAYEPLANNMDMSYYGEISIGTPPQNFLVLFDTGSSNLWVPSTLCQSQACANHN-- 113

```

Qy 127 DTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGFN-----TS 172
 || |||: :: ::| || || | ||| :| : ||
 Db 114 --EFDPNESSTFSTQDEFFSLQYSGSLTGIFGFDTV TI-QGISITNQEFGLSETEPGTS 170

Qy 173 FLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPN--VFSM 230
 || : ::||| |||: :: | : | : | : |||
 Db 171 FLYS-----PFDGILGLAFPSI---SAGGATTVMQKMLQENLLDFPVFSF 212

Qy 231 QMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNL 290
 : | | : || || ||::|: || | | :||: : | :|| | :|||
 Db 213 YLSGQ-----EGSQGGELVFGGVDPNLYTGQITWTPVTQTTYWQIGIEDFAVGGQSSGW 266

Qy 291 DCREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSET 350
 | : : |||: |||: || :| :|| :: : : | :| :| :| |
 Db 267 -CSQ--GCQGI VDTGTSLLTVPNQVFTELMQYIG-----AQADD---SGQYVASC SNIE- 314

Qy 351 PWSYFPKI-----SIYLRDENS---SRSEFRITILPQLYIQPMMGAGLNYECY 394
 | | | | | : || : || : | || ||:
 Db 315 ---YMPTITFVISGTSFPLPPSAYMLQSNSDYCTVGIESTYLPSQTGQPLW----- 362

Qy 395 RFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
 ::| : :| | :| :||| :
 Db 363 -----ILGDVFLRVYYSIYDMGNNQVGFATA 388

RESULT 6

A29937

gastricsin (EC 3.4.23.3) precursor - human

N;Alternate names: pepsin C; pepsinogen C

C;Species: Homo sapiens (man)

C;Date: 17-Oct-1988 #sequence_revision 17-Oct-1988 #text_change 31-Mar-2000

C;Accession: A29937; A31811; PX0028; I54213; A91125; A23458

R;Hayano, T.; Sogawa, K.; Ichihara, Y.; Fujii-Kuriyama, Y.; Takahashi, K.

J. Biol. Chem. 263, 1382-1385, 1988

A;Title: Primary structure of human pepsinogen C gene.

A;Reference number: A29937; MUID:88087276; PMID:3335549

A;Accession: A29937

A;Molecule type: DNA

A;Residues: 1-388 <HAY>

R;Taggart, R.T.; Cass, L.G.; Mohandas, T.K.; Derby, P.; Barr, P.J.; Pals, G.; Bell, G.I.

J. Biol. Chem. 264, 375-379, 1989

A;Title: Human pepsinogen C (progastricsin). Isolation of cDNA clones, localization to chromosome 6, and sequence homology with pepsinogen A.

A;Reference number: A31811; MUID:89079679; PMID:2909526

A;Accession: A31811

A;Molecule type: mRNA

A;Residues: 1-388 <TAG>

A;Cross-references: GB:J04443; NID:g551175; PIDN:AAA60074.1; PID:g551176

R;Athauda, S.B.P.; Tanji, M.; Kageyama, T.; Takahashi, K.

J. Biochem. 106, 920-927, 1989

A;Title: A comparative study on the NH2-terminal amino acid sequences and some other properties of six isozymic forms of human pepsinogens and pepsins.

A;Reference number: PX0023; MUID:90130402; PMID:2515193

A;Accession: PX0028

A;Molecule type: protein

A;Residues: 17-101 <ATH>

R;Pals, G.; Azuma, T.; Mohandas, T.K.; Bell, G.I.; Bacon, J.; Samloff, I.M.; Walz, D.A.; Barr, P.J.; Taggart, R.T.
 Genomics 4, 137-148, 1989
 A;Title: Human pepsinogen C (progastricsin) polymorphism: evidence for a single locus located at 6p21.1-pter.
 A;Reference number: I54213; MUID:89290840; PMID:2567697
 A;Accession: I54213
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-388 <RES>
 A;Cross-references: GB:M23077; NID:g189830; PIDN:AAA60063.1; PID:g387015; GB:J03063
 A;Note: parts of this sequence, including the amino end and carboxyl ends of the mature protein, were determined by protein sequencing
 R;Foltmann, B.; Jensen, A.L.
 Eur. J. Biochem. 128, 63-70, 1982
 A;Title: Human progastricsin. Analysis of intermediates during activation into gastricsin and determination of the amino acid sequence of the propeptide.
 A;Reference number: A91125; MUID:83079318; PMID:6816595
 A;Accession: A91125
 A;Molecule type: protein
 A;Residues: 17-39, 'ED', 42-51, 'S', 53-64 <FOL>
 A;Note: pro-form; 29-Leu was also found
 A;Note: activation at pH 2 is proposed to involve conformation change, cleavage after Phe-42, and cleavage after Leu-59
 C;Genetics:
 A;Gene: GDB:PGC
 A;Cross-references: GDB:119485; OMIM:169740
 A;Map position: 6p21.3-6p21.1
 A;Introns: 20/2; 70/3; 110/1; 149/3; 216/2; 256/2; 305/3; 338/3
 C;Superfamily: pepsin
 C;Keywords: aspartic proteinase; hydrolase; protein digestion; stomach; zymogen
 F;1-16/Domain: signal sequence #status predicted <SIG>
 F;17-59/Domain: propeptide #status experimental <PRO>
 F;60-388/Product: gastricsin #status experimental <MAT>

Query Match 13.1%; Score 353; DB 2; Length 388;
 Best Local Similarity 29.1%; Pred. No. 1.1e-18;
 Matches 120; Conservative 65; Mismatches 120; Indels 108; Gaps 21;

Qy	52	PAERHADG-LALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVD	110
		:: :: : : : : :	
Db	50	PAWKYRFGDLSVTYEP-----MAYMD-----AAYFGEISIGTPPQNFLVLF	91
Qy	111	TGSSNFAV-----AGTPHSYIDTYFDTERSSSTYRSKGFDTVKYTGQSWTGFVGEDL	162
		: : : ::	
Db	92	TGSSNLWVPSVYCQSQACTSHS----RFPNPSESSTYSTNGQTFSLQYGSGLTGFFGYDT	147
Qy	163	VTIPKGFNTSFLVNIATIFESENFLLPG-----IKWNGILGLAYATLAKPSSSLETFFDS	217
		: : : : : : :	
Db	148	LTV-----QSIQVPNQEFGLSEN--EPGTNFVYAQFDGIMGLAYPALSVDEAT--TAMQG	198
Qy	218	LVTQANIPN-VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQI	276
		: : : : : : : : : : : : :	
Db	199	MVQEGALTSPVFSVYLSNQ-----QGSSGGAUVFEGVDSSLYTGQIYWAPVTQELYWQI	252
Qy	277	EILKLEIGGQSLNLDREYNADKAIVDSGTTLRLPQKVFDVAVVEAVARASLIPEFSDGF	336

```

      | : ||||: | | :||||:||||: ||: |::|
Db      253 GIEEFLIGGQASGW-CSE--GCQAIVDGTGTSLLTVPQQYMSALLQA----- 295

Qy      337 WTGSQLACWTNSETPWSYF-----PKISIIYLRDENSRRSFRITILPQLYIQPMMG 386
      ||:| | : | | : : | |
Db      296 -TGAQ-----EDEYGQFLVNCNSIQNLPSLTFII-----NGVEFPLPPSSYI----- 336

Qy      387 AGLNYECY-RFGISP----STNA---LVIGATVMEGFYVIFDRAQKRVGFAAS 431
      |: | |: | | :| : :| :||| :
Db      337 --LSNNGYCTVGVEPTYLSSQNGQPLWILGDVFLRSYYSVYDLGNNRVGFATA 387

```

RESULT 7

JC7246

pepsinogen C - common marmoset

C;Species: Callithrix jacchus (common marmoset)

C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000

C;Accession: JC7246

R;Kageyama, T.

J. Biochem. 127, 761-770, 2000

A;Title: New world monkey pepsinogens A and C, and prochymosins. Purification, characterization of enzymatic properties, cDNA cloning, and molecular evolution.

A;Reference number: JC7245

A;Accession: JC7246

A;Molecule type: mRNA

A;Residues: 1-388 <KAG>

A;Cross-references: DDBJ:AB038385

A;Experimental source: strain NW791

C;Comment: This protein, a zymogen of pepsins, is the major proteolytic enzyme in vertebrate gastric juices. It plays roles in gastric digestion, and is a useful molecular marker for clarifying the evolution of mammalian orders and families.

C;Superfamily: pepsin

C;Keywords: gastric juice; zymogen

```

Query Match          13.1%;  Score 351.5;  DB 2;  Length 388;
Best Local Similarity 30.1%;  Pred. No. 1.4e-18;
Matches 112;  Conservative 56;  Mismatches 115;  Indels 89;  Gaps 17;

```

```

Qy      92 YYLEMLIGTPPQKLQILVDTGSSNFAV-----AGTPHSYIDTYFDTERSSTYRSKGF 143
      |: |: ||||| | :| ||||| | | | | | : | ||| | |
Db      73 YFGEISIGTPPQNFLVLFDTGSSNLWVPSVYCQSQACTSHS----RFNPSASSTYSSNGQ 128

Qy     144 DVTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATIFESENFFLPG-----IKWNGILG 198
      :::| || ||| | | :|: | | | | | | | | | | | | | | | | | | | | |
Db     129 TFSLQYGSGLTGFFGYDTLTV-----QSIQVPNQEFGLSEN--EPGTNFVYAQFDGIMG 181

Qy     199 LAYATLAKPSSSLETFFDSLVTQANIPN-VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPS 257
      ||| |: : : | : : : : | | | | | | | | | | | | | | | | | | | | |
Db     182 LAYPALSMGGAT--TAMQGMLOEGALTSPVFSFYLSNQ-----QGSSGGAVIFGGVDSS 233

Qy     258 LYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKAIVDSGTTLRLPQKVFD 317
      || | |: : |: :| |:|| | : ||||: | | :||||:||||: ||:
Db     234 LYTGQIYWAPVTQELYWQIGIEEFLIGGQASGW-CSE--GCQAIVDGTGTSLLTVPQQYMS 290

Qy     318 AVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYF-----PKISIIYLRDENS 367
      | :|| | | | | | | : | | : :

```

```

Db      291 AFLEA-----TGAQ-----EDEYGQFLVNCDSIQNLPTLTFII----- 323
Qy      368 SRSFRITILPQLYIQPMMGAGLNYECY-RFGISP-----STNALVIGATVMEGFYVIF 419
          : | || | : | | : | | : | : | : |
Db      324 -NGVEFPLPPSSYI-----LSNNGYCTVGVEPTYLSSQNSQPLWILGDVFLRSYYSVF 375
Qy      420 DRAQKRVGFAAS 431
          | |||| :
Db      376 DLGNNRVGFATA 387

```

RESULT 8

B43356

gastricsin (EC 3.4.23.3) precursor - guinea pig

N;Alternate names: pepsin C

C;Species: Cavia porcellus (guinea pig)

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999

C;Accession: B43356

R;Kageyama, T.; Ichinose, M.; Tsukada, S.; Miki, K.; Kurokawa, K.; Koiwai, O.;
TANJI, M.; Yakabe, E.; Athauda, S.B.; Takahashi, K.
J. Biol. Chem. 267, 16450-16459, 1992

A;Title: Gastric procathepsin E and progastricsin from guinea pig. Purification,
molecular cloning of cDNAs, and characterization of enzymatic properties, with
special reference to procathepsin E.

A;Reference number: A43356; MUID:92355614; PMID:1644829

A;Accession: B43356

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-394 <KAG>

A;Cross-references: GB:M88652; NID:g191296; PIDN:AAA37053.1; PID:g191297

A;Note: sequence extracted from NCBI backbone (NCBIN:110805, NCBIP:110806)

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion;
stomach

```

Query Match          12.1%; Score 324.5; DB 2; Length 394;
Best Local Similarity 29.0%; Pred. No. 1.5e-16;
Matches 107; Conservative 63; Mismatches 116; Indels 83; Gaps 18;

```

```

Qy      92 YYLEMLIGTPPQKLQILVDTGSSNF-----AVAGTPHSYIDTYFDTERSSTYRSKGF 143
          |: :: :| || | | :| || || | : | | : | | | :
Db      79 YFGQISLGTTPPQSFQVLFDTGSSNLWVPSVYCSSLACTTH----TRFNPRDSSTYVATDQ 134
Qy      144 DVTVKYTQGSWTGFVGEDLVTI-----PK-GFNTSFLVNIATIFESENFFLPG-----IK 192
          ::| | | | | | :| | | | | | | | | | | | :
Db      135 SFSLEYGTGSLTGFGYDTMTIQDIQVPKQEFGLS-----ETE----PGSDFVYAE 181
Qy      193 WNGILGLAYATLAKPSSSLETFFDLSVTQANI-PNVFSMQMCGAGLPVAGS--GTNGGSL 249
          ::||| | | | : : | | : : : : | | | | : | |
Db      182 FDGILGLGYPLSEGGAT--TAMQGLLREGALSQSLFSVYL-----GSQQGSDEGQL 231
Qy      250 VLGgiePSLYKGDWYTPiKEEWYYQIEILKLEIGGQSLNLDcreYNADKAIVDSGTtLL 309
          :|||: : || | || :|| : | | : | | : | | :||: ||
Db      232 ILGGVDESlytGDIYwTPvtQELyWQIGIEGFLIDGSASGWCSR--GCQGIVDTGTsLL 288
Qy      310 RLPQKVFDaVVEaVaRaSLiPEfSDGfWtGSQLaCWtNSeTPwSYfPKiSiYlRDeNsSR 369
          :| :| :| :| :| :| :| :| :| :| :| :| :| :

```

```

Db      289 TVPSDYLSTLVQAIGAE--NEYGEYF-----VSCSSIQDLPTLTFVISGV----- 332
Qy      370 SFRITILPQLYIQP-----MMGAGLNYECYRFGISPSTN--ALVIGATVMEGFYVIFDRA 422
          : | ||      | : |      |      : ||      : : |      : : | : : | |
Db      333 --EFPLSPSAYILSGENYCMVGLESTY-----VSPGGGEPVWILGDVFLRSYYSVYDLA 384
Qy      423 QKRVGFAAS 431
          ||||| :
Db      385 NNRVGFATA 393

```

RESULT 9

JC7575

pepsinogen A - bullfrog

C;Species: Rana catesbeiana (bullfrog)

C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001

C;Accession: JC7575

R;Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.

J. Biochem. 129, 147-153, 2001

A;Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens, and molecular cloning of Xenopus and bullfrog pepsinogens.

A;Reference number: JC7573; MUID:21064922; PMID:11134969

A;Contents: Stomach

A;Accession: JC7575

A;Molecule type: mRNA

A;Residues: 1-385 <IKU>

A;Cross-references: DDBJ:AB045376

C;Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like activity.

C;Genetics:

A;Gene: PgA

C;Superfamily: pepsin

C;Keywords: stomach; zymogen

```

Query Match          11.9%;  Score 320;  DB 2;  Length 385;
Best Local Similarity 27.8%;  Pred. No. 3.2e-16;
Matches 111;  Conservative 67;  Mismatches 147;  Indels 74;  Gaps 15;

```

```

Qy      50 GTPAERHADGLALALEPALASPAGAANFLAMVDNLQDSGRGYYLEMLIGTPPQKLQILV 109
          | : : |      |      | : || : |      : ||      | : : ||||| : :
Db      39 GDYLKKHHYNPATKYFPSLAQASG-----EPLQNYMDIEYFGTISIGTPPQSFTVIF 90
Qy      110 DTGSSNFAVAGTPHSYIDT-----YFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDL 162
          ||||| |      | : :      | : : ||| : : :      | : : | || : || : |
Db      91 DTGSSNLWV---PSVYCSPACTNHHMFNPQQSSTFQATNTPVSIQYGTGMSGFLGYDT 147
Qy      163 VTIPKGFNTSFLVNIATIFESE-NFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQ 221
          | :      |      | : : || ||      : : ||||| : : || ||      || : : |
Db      148 VQVG--NIQITNQIFGLSQSEPGSFLYYSFPDGILGLAFPSLA--SSQATPVFDNMWNQ 202
Qy      222 ANIP-NVFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILK 280
          || : : || : :      : | : | : : || : : | | : : : | : | | : || :
Db      203 GLIPQDLFSVYL-----SSQGQSGSFVLFGGVDTSYYTGNLNWVPLTAETYWQITVDS 255
Qy      281 LEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGS 340
          : |||| :      : |||| : || : || |      | : :      | : : | : | :
Db      256 ISIGGQVIACS----GSCSAIVDTGTSLLAGPSTPI-ANIQYYIGAN---QDSNGQYV-- 305

```

```

Qy      341 QLACWTNSETPWSYFP-----KISIIYLRDENSS--RSFRITILPQLYIQPMMGAGLN 390
          : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      306 -INCNNISNMPTVVFTINGVQYPLPASAYVRQSQQSCTSGFQAMNLP----- 351

Qy      391 YECYREGISPSTNALVIGATVMEGFYVIFDRAQKRVGFA 429
          : | : : : | : : | | | | | | | | | | | | | | | |
Db      352 -----TSSGDLWILGDVFIREFYYVVFDRANNYVAMA 382

```

RESULT 10

REMSK

renin (EC 3.4.23.15) precursor, renal - mouse

N;Alternate names: angiotensin-forming enzyme; angiotensinogenase; renin 1

C;Species: Mus musculus (house mouse)

C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 18-Jun-1999

C;Accession: A00989; S07636; A22766; A22058; I57576; A05137; JH0083

R;Holm, I.; Ollo, R.; Panthier, J.J.; Rougeon, F.

EMBO J. 3, 557-562, 1984

A;Title: Evolution of aspartyl proteases by gene duplication: the mouse renin gene is organized in two homologous clusters of four exons.

A;Reference number: A00989; MUID:84182525; PMID:6370686

A;Accession: A00989

A;Molecule type: DNA

A;Residues: 1-402 <HOL>

A;Cross-references: EMBL:X00850

R;Kim, W.S.; Murakami, K.; Nakayama, K.

Nucleic Acids Res. 17, 9480, 1989

A;Title: Nucleotide sequence of a cDNA coding for mouse Ren1 preprorenin.

A;Reference number: S07636; MUID:90067953; PMID:2685761

A;Accession: S07636

A;Molecule type: mRNA

A;Residues: 1-402 <KIM>

A;Cross-references: EMBL:X16642; NID:g53930; PIDN:CAA34636.1; PID:g53931

R;Mullins, J.J.; Burt, D.W.; Windass, J.D.; McTurk, P.; George, H.; Brammar, W.J.

EMBO J. 1, 1461-1466, 1982

A;Title: Molecular cloning of two distinct renin genes from the DBA/2 mouse.

A;Reference number: A90968; MUID:84207899; PMID:6327270

A;Accession: A22766

A;Molecule type: mRNA

A;Residues: 269-314, 'D', 316 <MUL>

R;Panthier, J.J.; Dreyfus, M.; Roux, D.T.L.; Rougeon, F.

Proc. Natl. Acad. Sci. U.S.A. 81, 5489-5493, 1984

A;Title: Mouse kidney and submaxillary gland renin genes differ in their 5' putative regulatory sequences.

A;Reference number: A22058; MUID:84298161; PMID:6089205

A;Accession: A22058

A;Molecule type: DNA

A;Residues: 1-30 <PAN>

R;Field, L.J.; Philbrick, W.M.; Howles, P.N.; Dickinson, D.P.; McGowan, R.A.; Gross, K.W.

Mol. Cell. Biol. 4, 2321-2331, 1984

A;Title: Expression of tissue-specific Ren-1 and Ren-2 genes of mice: Comparative analysis of 5'-proximal flanking regions.

A;Reference number: I57576; MUID:85085936; PMID:6392850

A;Accession: I57576

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-31 <RES>
A;Cross-references: GB:K02800; NID:g200689; PIDN:AAA40044.1; PID:g200690
C;Comment: The only known function of renal renin is to release angiotensin I from angiotensinogen in the plasma, initiating a cascade of reactions that produces an elevation of blood pressure and increased sodium retention by the kidney.
C;Comment: Renal renin is synthesized by the juxtaglomerular cells of the kidney in response to decreased blood pressure and sodium concentration.
C;Genetics:
A;Gene: Ren-1
A;Introns: 31/2; 81/3; 123/1; 162/3; 228/2; 268/2; 316/3; 349/3
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; blood pressure control; glycoprotein; hydrolase; kidney; plasma
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-64/Domain: propeptide #status predicted <PRO>
F;65-402/Product: renin #status predicted <MAT>
F;69,139,320/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;102,287/Active site: Asp #status predicted

Query Match 11.9%; Score 320; DB 1; Length 402;
Best Local Similarity 28.6%; Pred. No. 3.4e-16;
Matches 126; Conservative 66; Mismatches 181; Indels 68; Gaps 21;

Qy	10	LPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPG-PGTPAERHADGLALALE---	65
		: : : : :	
Db	6	MPLWALLLL-----WSPCTFSLPTRTATFERIPLKKMPSVREILEERGVDMLRLSAEWGV	60
Qy	66	----PA---LASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAV	118
		: : : :: :	
Db	61	FTKRPSLTNLTSPVVLTNYL----NTQ-----YYGEIGIGTPPQTFKVIFDTGSANLWV	110
Qy	119	AGTPHSY-----IDTYFDTERSSTYRSKGFVDTVKYTQGSWTGFVGEDLVITPKGENTS	172
		: :: : : : : : :	
Db	111	PSTKCSRLYLACGIHSLYESSDSSSYMENGSDFTIHYGSGRVKGFLSQDSVTV-GGITVT	169
Qy	173	FLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANI-PNVFSMQ	231
		: : : : :: : :	
Db	170	QTFGEVTELPLIPFML--AKFDGVLGMGFP--AQAVGGVTPVFDHILSQGVLKEEVFSVY	225
Qy	232	MCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD	291
		: : : : : : : : :	
Db	226	Y-----NRGSHLLGGEVVLGGSDPQHYQGNGFHYVSISKTDWQITMKGVSVG--SSTLL	277
Qy	292	CREYNADKAIVDSGTTLLRLPQKVFDVAVVEAV-ARASLIPEFSDGFWTGSQACWTNSET	350
		: : : : : : : : : :	
Db	278	CEEGCA--VVVDTGSSFISAPTSSLKLIMQALGAKEKRIEY-----VVNC---SQV	324
Qy	351	PWSYFPKISIIYLRDENSSRSFRITILPQLYIQPMMGAGL-NYECYRFGISPSTNAL-VIG	408
		: : : : : : :	
Db	325	P--TLPDISFDL----GGRAYTLSSDYVLQYPNRRDKLCTLALHAMDIPPTGPVWVLG	378
Qy	409	ATVMEGFYVIFDRAQKRVGFA	429
		: :	
Db	379	ATFIRKFYTEFDRHNNRIGFA	399

RESULT 11

S66516

oryzasin (EC 3.4.23.-) precursor - rice

N;Alternate names: aspartic proteinase 1

C;Species: Oryza sativa (rice)

C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000

C;Accession: S66516; S66517

R;Asakura, T.; Watanabe, H.; Abe, K.; Arai, S.

Eur. J. Biochem. 232, 77-83, 1995

A;Title: Rice aspartic proteinase, oryzasin, expressed during seed ripening and germination, has a gene organization distinct from those of animal and microbial aspartic proteinases.

A;Reference number: S66516; MUID:96048031; PMID:7556174

A;Accession: S66516

A;Molecule type: DNA

A;Residues: 1-509 <ASA>

A;Cross-references: EMBL:D32165; NID:g511665; PIDN:BAA06876.1; PID:g1030715

A;Accession: S66517

A;Molecule type: mRNA

A;Residues: 1-509 <ASZ>

A;Cross-references: EMBL:D32144; NID:g1255684; PIDN:BAA06875.1; PID:g1711289

C;Comment: The pair of saposin repeat homology domains tagged SAP1 and SAP2 represent a cyclical permutation of a single saposin repeat.

C;Genetics:

A;Introns: 119/3; 140/1; 171/3; 209/2; 265/3; 279/1; 300/3; 338/3; 360/2; 412/3; 452/3; 482/2

C;Superfamily: oryzasin; saposin repeat homology

C;Keywords: aspartic proteinase; hydrolase

F;1-20/Domain: signal sequence #status predicted <SIG>

F;21-68/Domain: propeptide #status predicted <PRO>

F;68-509/Product: aspartic proteinase 1 #status predicted <MAT>

F;316-361/Domain: saposin repeat homology #status atypical <SAP1>

F;370-420/Domain: saposin repeat homology #status atypical <SAP2>

F;103,290/Active site: Asp #status predicted

Query Match 11.7%; Score 313.5; DB 2; Length 509;

Best Local Similarity 23.0%; Pred. No. 1.5e-15;

Matches 127; Conservative 75; Mismatches 179; Indels 171; Gaps 19;

```

Qy      3 ALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL 62
        ::|  ||  :|  ||  |:  |      |  |      :||  |  |      |  |  |
Db      5 SVALVLLAAVLLQALLPASAEGLVRIALKKRPIDENSRVAARLSG-----EEGARRLGL 59

Qy     63 ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSN----- 115
        :|      |  :  :|:  :  :      |:  |:  :|||  ||  ::  |||  ||
Db     60 RGANSLGGGGGEGDIVALKNYMNAQ----YFGEIGVGTPPQKFTVIFDTGSSNLWVPSAK 115

Qy    116 --FAVAGTPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSF 173
        |::|  ||      :  :  :|||  |:  |      :|  ||  ||  ||  ||:
Db    116 CYFSIACFFHS----RYKSGQSSTYQKNGKPAAIQYGTGSIAGFFSEDSVTVGD----- 165

Qy    174 LVNIATIFESENFF----LPGI-----KWNIGILGLAYATLAKPSSSLETFFDSLVTQANI 224
        :  :  :  |      ||:      |::|||  |  :  :  :      :
Db    166 -----LVVKDQEFIEATKEPGLTFMVAKFDGILGLGFQEISVGDA-----V 206

```

Qy 225 PNVFSMQMCG-AGLPVAGSGTN-----GGSLVLGGIEPSLYKGDIWYTPIKEEWYYQI 276
| : | | | | | | | | : | | : | | : | | : | : | : |
Db 207 PVWYKMVEQGLVSEPVFSFWFNHRSDEGE GGEIVFGGMDPSHYKGNHTYVPVSQKGYWQF 266

Qy 277 EILKLEIGGQSLNLD CREYNADKAI VDSGTTLLRLPQKVFD AVEAVARASLIPE----- 331
| : : | | : : | : | | | | : | : : : : : : :
Db 267 EMGDVLIGGKTTGF-CA--SGCSAIADSGTSLLAGPTAIITEINEKIGATGVVSQECKTV 323

Qy 332 -----FSDGF----- 336
| : |
Db 324 VSQYGQQILDLLLAETQPSKICSQVGLCTFDGKHGVSAGIKSVVDDEAGESNGLQSGPMC 383

Qy 337 -----WTGSQLACWTNSETPWSY-----FPKISIIYLRD 364
| : | | : : | | : | : | : | :
Db 384 NACEMAVVWMQNQLAQNKTDLILNYINQLCDKLPSPMGESSVDCGSLASMPEISFTIGA 443

Qy 365 ENSSRSFRITILPQLYIQPMMGAGLNYECY----RFGISPSTNAL-VIGATVMEGFYVIF 419
: : : | : | : | : | : | : | : : : |
Db 444 K-----KFALKPEEYIL-KVGEGAAAQCISGFTAMDIPPPRGPLWILGDVFMGAYHTVF 496

Qy 420 DRAQKRVGFAAS 431
| : | | | | |
Db 497 DYGKM RVGFAKS 508

RESULT 12

A24608

gastricsin (EC 3.4.23.3) precursor - rat

N;Alternate names: pepsinogen C

N;Contains: pepsin A (EC 3.4.23.1) precursor

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Jun-1988 #sequence_revision 05-Aug-1994 #text_change 18-Jun-1999

C;Accession: A33510; A24608; C22434; A05145; A61298

R;Ishihara, T.; Ichihara, Y.; Hayano, T.; Katsura, I.; Sogawa, K.; Fujii-Kuriyama, Y.; Takahashi, K.

J. Biol. Chem. 264, 10193-10199, 1989

A;Title: Primary structure and transcriptional regulation of rat pepsinogen C gene.

A;Reference number: A33510; MUID:89255508; PMID:2722863

A;Accession: A33510

A;Molecule type: DNA

A;Residues: 1-392 <ISH>

A;Cross-references: GB:M25985

R;Ichihara, Y.; Sogawa, K.; Morohashi, K.; Fujii-Kuriyama, Y.; Takahashi, K. Eur. J. Biochem. 161, 7-12, 1986

A;Title: Nucleotide sequence of a nearly full-length cDNA coding for pepsinogen of rat gastric mucosa.

A;Reference number: A24608; MUID:87054020; PMID:3780741

A;Accession: A24608

A;Molecule type: mRNA

A;Residues: 1-392 <ICH>

A;Cross-references: GB:X04644; NID:g56880; PIDN:CAA28305.1; PID:g56881

R;Ichihara, Y.; Sogawa, K.; Takahashi, K.

J. Biochem. 98, 483-492, 1985

A;Title: Isolation of human, swine, and rat prepepsinogens and calf preprochymosin, and determination of the primary structures of their NH2-terminal signal sequences.

A;Reference number: A22434; MUID:86059312; PMID:2415509
A;Accession: C22434
A;Molecule type: protein
A;Residues: 1-19,'X',21-23,'X',25-29 <IC2>
R;Arai, K.M.; Muto, N.; Tani, S.; Akahane, K.
Biochim. Biophys. Acta 788, 256-261, 1984
A;Title: The N-terminal sequence of rat pepsinogen.
A;Reference number: A05145; MUID:84257697; PMID:6743670
A;Accession: A05145
A;Molecule type: protein
A;Residues: 17-30,'Q',32-102,'A',104-108,'L',110-112 <ARA>
A;Experimental source: Wistar strain
R;Ichihara, Y.; Sogawa, K.; Takahashi, K.
J. Biochem. 92, 603-606, 1982
A;Title: Rat gastric prepepsinogen: in vitro synthesis and partial amino-terminal signal sequence.
A;Reference number: A61298; MUID:83030750; PMID:6182139
A;Accession: A61298
A;Molecule type: protein
A;Residues: 1,'XX',4-6,'X',8-9,'X',11,'X',13-14,'XXX',18-19,'X',21,'X',23,'XX',26,'X' <IC3>
C;Comment: This enzyme has more restricted specificity than pepsin A. It is the major form of pepsinogen in rat gastric mucosa.
C;Genetics:
A;Introns: 20/2; 73/3; 113/1; 152/3; 219/2; 259/2; 309/3; 342/3
A;Note: there are at least two very similar genes for gastricsin in rat
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach
F;1-16/Domain: signal sequence #status experimental <SIG>
F;17-392/Product: pepsinogen #status experimental <MAT>
F;17-62/Domain: activation peptide #status experimental <ACT>
F;94,280/Active site: Asp #status predicted
F;107-112,270-275,314-347/Disulfide bonds: #status predicted

Query Match 11.6%; Score 313; DB 1; Length 392;
Best Local Similarity 29.5%; Pred. No. 1.1e-15;
Matches 105; Conservative 56; Mismatches 139; Indels 56; Gaps 16;

Qy	92	YYLEMLIGTPPQKLQILVDTGSSNFAV-----AGTPHSYIDTYFDTERSSTYRSKGF	143
		: : : : : :	
Db	76	YFGEISIGTPPQNFLVLFDTGSSNLWVSSVYCQSEACTTHA----RFNPSKSSTYYTEGQ	131
Qy	144	DVTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATIFESENFPLPG-----IKWNGILG	198
		::: : : :: :	
Db	132	TESLQYGTGSLTGFFGYDTLTV-----QSIQVPNQEFGLSEN--EPGTNFVYAQFDGIMG	184
Qy	199	LAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGS--GTNGGSLVLGGIEP	256
		: : : : : : : ::	
Db	185	LAYPGLS--SGGATTALQGM LGE----GALSQPLFGVYL---GSQQGSNGGQIVFGGVDK	235
Qy	257	SLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKAIVDSGTTLLRLPQKVF	316
		: : : : : : : : : : : :	
Db	236	NLYTGEITWVPVTQELYWQITIDDFLIGDQASGW-CSSQGC-QGIVDTGTSLLVMPAQYL	293
Qy	317	DAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISIIYLRDENS SRFRITIL	376
		::: : : : :: : : : :	

```

Db          294 SELLQTIGAQE--GEYGEYF-----VSCDSVSS-----LPTLSFVL-----NGVQFPLS 335
Qy          377 PQLY-IQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
              | | ||      : |              ::| : :| ||| :|| | |
Db          336 PSSYIIQEDNFCMVGLESLTSESGQPLWILGDVFLRSYYAIFDMGNNKVGLATS 391

```

RESULT 13

A41443

pepsin (EC 3.4.23.-) precursor, embryonic - chicken

C;Species: Gallus gallus (chicken)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 21-Jul-2000

C;Accession: A41443

R;Hayashi, K.; Agata, K.; Mochii, M.; Yasugi, S.; Eguchi, G.; Mizuno, T.

J. Biochem. 103, 290-296, 1988

A;Title: Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken pepsinogen: phylogenetic relationship with prochymosin.

A;Reference number: A41443; MUID:88227903; PMID:3131317

A;Accession: A41443

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-383 <HAY>

A;Cross-references: GB:D00215; NID:g2760810; PIDN:BAA00153.1; PID:g222853

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; hydrolase; protein digestion

```

Query Match          11.5%; Score 310; DB 2; Length 383;
Best Local Similarity 26.8%; Pred. No. 1.8e-15;
Matches 106; Conservative 63; Mismatches 136; Indels 90; Gaps 15;

```

```

Qy          56 HA--DGLALALEPALASPAGAAFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGS 113
              || | | : || | | | | | | | | | | : ||||| : : ||||
Db          55 HAFPDVLTVVTEPLL-----NTLDM-----EYYGTISIGTPPQDFTVVFDTG 97

Qy          114 SNFAVAG----TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGF 169
              || | : | : : ||||| : | : : : | | | | || | || :
Db          98 SNLWVPSVSCTSPACQSHQMFNPSQSSTYKSTGQNLSIHGTGDMEGTVGCDTVTVASLM 157

Qy          170 NTSFLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANI-PNVF 228
              : : | : : | | || : : : ||||| | : || : : || : : | : |
Db          158 DTNQLFGLST-SEPGQFFV-YVKFDGILGLGYPSLA--ADGITPVFDNMVNESLLEQNLF 213

Qy          229 SMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSL 288
              | : : : | : | ||| : | : | | : | : : | : | : : : | :
Db          214 SVYLSREPM-----GSMVVFGGIDESYFTGSINWIPVSYQGYWQISMDSIIVNKQEI 265

Qy          289 NLDCREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNS 348
              : : ||| : ||| : | : : : || |
Db          266 ACS----SGCQAIIDTGTSLVAGPASDINDIQSAVG-----ANQ 300

Qy          349 ETPWSYFPKISIIYLRDENSSRSFRITILPQLYIQPMGAGLNY-----ECY 394
              | | | | | | | : | : : : | : | | |
Db          301 NTYGEY-----SVNCSHILAMPDVVF--VIG-GIQYPVPALAYTEQNGQGTCM 345

Qy          395 RFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFA 429
              : | : : : | : : | |||| | || |
Db          346 SSFQNSSADLWILGDVFIRVYYSIFDRANNRVGLA 380

```

RESULT 14

KHHUD

cathepsin D (EC 3.4.23.5) precursor [validated] - human

N;Alternate names: preprocathepsin D

C;Species: Homo sapiens (man)

C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 15-Sep-2000

C;Accession: A25771; S30749; PC2066; I59236; I57716

R;Faust, P.L.; Kornfeld, S.; Chirgwin, J.M.

Proc. Natl. Acad. Sci. U.S.A. 82, 4910-4914, 1985

A;Title: Cloning and sequence analysis of cDNA for human cathepsin D.

A;Reference number: A25771; MUID:85270436; PMID:3927292

A;Accession: A25771

A;Molecule type: mRNA

A;Residues: 1-412 <FAU>

A;Cross-references: EMBL:M11233; NID:g181179; PIDN:AAB59529.1; PID:g181180

R;Westley, B.R.; May, F.E.B.

Nucleic Acids Res. 15, 3773-3786, 1987

A;Title: Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive human breast cancer cells.

A;Reference number: S30749; MUID:87231068; PMID:3588310

A;Accession: S30749

A;Molecule type: mRNA

A;Residues: 1-412 <WES>

A;Cross-references: EMBL:X05344; NID:g29677; PIDN:CAA28955.1; PID:g29678

R;May, F.E.B.; Smith, D.J.; Westley, B.R.

Gene 134, 277-282, 1993

A;Title: The human cathepsin D-encoding gene is transcribed from an estrogen-regulated and a constitutive start point.

A;Reference number: PC2066; MUID:94085791; PMID:8262386

A;Accession: PC2066

A;Molecule type: DNA

A;Residues: 1-23 <MAY>

A;Cross-references: GB:L12980; NID:g291930; PIDN:AAA16314.1; PID:g455429

A;Experimental source: MCF-7 cell

R;Cavaillès, V.; Augereau, P.; Rochefort, H.

Proc. Natl. Acad. Sci. U.S.A. 90, 203-207, 1993

A;Title: Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate only TATA-dependent transcription in breast cancer cells.

A;Reference number: I59236; MUID:93126342; PMID:8419924

A;Accession: I59236

A;Status: translation not shown; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-22 <CAV1>

A;Cross-references: GB:S52557; NID:g263124; PIDN:AAD13868.1; PID:g4261568

R;Augereau, P.; Miralles, F.; Cavaillès, V.; Gaudelot, C.; Parker, M.; Rochefort, H.

Mol. Endocrinol. 8, 693-703, 1994

A;Title: Characterization of the proximal estrogen-responsive element of human cathepsin D gene.

A;Reference number: I57716; MUID:95021301; PMID:7935485

A;Accession: I57716

A;Status: translation not shown; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-22 <CAV2>

A;Cross-references: GB:S74689; NID:g786350; PIDN:AAD14156.1; PID:g4261856

R;Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A;Reference number: A51839; PDB:1LYA
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 65-161;170-241
R;Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A;Reference number: A51840; PDB:1LYB
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, with inhibitor residues 65-161;170-241
R;Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Hosur, M.V.; Sowder II, R.C.; Cachau, R.E.; Collins, J.; Silva, A.M.; Erickson, J.W.
Proc. Natl. Acad. Sci. U.S.A. 90, 6796-6800, 1993
A;Title: Crystal structures of native and inhibited forms of human cathepsin D: implications for lysosomal targeting and drug design.
A;Reference number: A48229; MUID:93342076; PMID:8393577
A;Contents: annotation; X-ray crystallography, 2.5 angstroms
C;Comment: Cathepsin D is a ubiquitous lysosomal proteinase.
C;Comment: In addition to the propeptide, residues 163-168 and 411-412 are proteolytically removed. Residues 169 and 170 are also partially removed.
C;Comment: The carbohydrate bound to 134-Asn contains a mannose-6-phosphate that is bound near 267-Lys and the phosphotransferase recognition region.
C;Genetics:
A;Gene: GDB:CTSD
A;Cross-references: GDB:120512; OMIM:116840
A;Map position: 11p15.5-11p15.5
C;Function:
A;Description: limited specificity endopeptidase
A;Pathway: intracellular protein degradation
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-64/Domain: propeptide #status predicted <PRO>
F;65-162,169-410/Product: cathepsin D #status experimental <MAT>
F;267,329-356/Region: phosphotransferase recognition
F;91-160,110-117,286-290,329-366/Disulfide bonds: #status experimental
F;97,295/Active site: Asp #status experimental
F;134,263/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 11.5%; Score 308.5; DB 1; Length 412;
Best Local Similarity 27.1%; Pred. No. 2.5e-15;
Matches 121; Conservative 75; Mismatches 180; Indels 71; Gaps 22;

```

Qy      9 LLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPG-----PGTPAERHADGLAL 62
          ||||      | |||  | |  :||      :  | ::  |      |  ::  :
Db      6 LLPLAL--CLLAAP--ASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPVSKYSQAVPA 61

Qy     63 ALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTP 122
          |  :  |      |::      || | : |||||  :: |||||  |
Db     62 VTEGPI--PEVLKNYM-----DAQYYGEIGIGTPPQCFTVVFDTGSSNLWVPSIH 109

Qy    123 HSYIDT-----YFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIP--KGFNTSFL 174
          :|      ::::|||  |      :  |  || :|:: :| |::|  :  | |
Db    110 CKLLDIACWIHHKYNSDKSSTYVKNGTSFDIHYGSGSLSGYLSQDTSVSVPCQSASSASAL 169

Qy    175 --VNIATIFESENFFLPGI-----KWNGILGLAYATLAKPSSSLETFFDSLVTQANI-PN 226

```

```

      | :      |    |||      |::|||:|  ::   ::   ||::|  | :  |
Db      170 GGVKVERQVFGEATKQPGITFIAAKFDGILGMAYPRIS--VNNVLPVFDNLMQQKLVDQN 227

Qy      227 VFQMCMGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQ 286
      :||      |      || |::|  :   ||| : | : : |:: : ::|:
Db      228 IFSFY-----LSRDPDAQPGGELMLGGTDSKYYKGSLSYLVNTRKAYWQVHLDQVEV-AS 281

Qy      287 SLNLDCREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWT 346
      | | |::| :|||:|::|  |   | | |      :|      :      : |
Db      282 GLTL-CKE--GCEAIVDTGTSLMVG---VDEVRELQKAIGAVPLIQGEY---MIPC-- 329

Qy      347 NSETPWSYFPKISIIYLRDENSRSFRITILPOLYIQPMGAGLNYECYRF---GISPSTN 403
      | | |::| : :::: |: | : | |      |   | | :
Db      330 ---EKVSTLPAITLKL----GGKGYKLS--PEDYTLKVSQAGKTLCLSGFMGMDIPPPSG 380

Qy      404 AL-VIGATVMEGFYVIFDRAQKRVGFA 429
      | ::| : : | :||  ||||
Db      381 PLWILGDVFIGRYTTFDRDNNRVGFA 407

```

RESULT 15

KHMSD

cathepsin D (EC 3.4.23.5) precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999

C;Accession: I48278; S14704; S12587

R;Hetman, M.; Perschl, A.; Saftig, P.; Von Figura, K.; Peters, C.

DNA Cell Biol. 13, 419-427, 1994

A;Title: Mouse cathepsin D gene: molecular organization, characterization of the promoter, and chromosomal localization.

A;Reference number: I48278; MUID:94280622; PMID:8011168

A;Accession: I48278

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-410 <RES>

A;Cross-references: EMBL:X68378; NID:g50302; PIDN:CAA48453.1; PID:g817945

R;Diedrich, J.F.; Staskus, K.A.; Retzel, E.F.; Haase, A.T.

Nucleic Acids Res. 18, 7184, 1990

A;Title: Nucleotide sequence of a cDNA encoding mouse cathepsin D.

A;Reference number: S14704; MUID:91088345; PMID:2263503

A;Accession: S14704

A;Molecule type: mRNA

A;Residues: 1-410 <DIE>

A;Cross-references: EMBL:X53337; NID:g50300; PIDN:CAA37423.1; PID:g50301

R;Grusby, M.J.; Mitchell, S.C.; Glimcher, L.H.

Nucleic Acids Res. 18, 4008, 1990

A;Title: Molecular cloning of mouse cathepsin D.

A;Reference number: S12587; MUID:90326544; PMID:2374732

A;Accession: S12587

A;Molecule type: mRNA

A;Residues: 1-410 <GRU>

A;Cross-references: EMBL:X52886; NID:g50298; PIDN:CAA37067.1; PID:g50299

C;Genetics:

A;Introns: 23/2; 76/3; 118/1; 157/3; 233/2; 274/2; 322/3; 355/3

C;Function:

A;Description: limited specificity endopeptidase

A;Pathway: intracellular protein degradation

C;Superfamily: pepsin
 C;Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation
 F;1-20/Domain: signal sequence #status predicted <SIG>
 F;21-64/Domain: propeptide #status predicted <PRO>
 F;65-410/Product: cathepsin D, single-chain form #status predicted <MAT>
 F;91-160,110-117,284-288,327-364/Disulfide bonds: #status predicted
 F;97,293/Active site: Asp #status predicted
 F;134,261/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.4%; Score 306.5; DB 1; Length 410;
 Best Local Similarity 27.5%; Pred. No. 3.6e-15;
 Matches 103; Conservative 64; Mismatches 123; Indels 85; Gaps 15;

```

Qy      92 YYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDT-----YFDTERSSTYRSKGFDV 145
      || :: |||||  :: |||||  |      :|      ::::| |||  |
Db      79 YYGDIGIGTPPQCFTVVFDTGSSNLWVPSIHCKILDIAWVHHKYNSDKSSTYVKNGTSF 138

Qy     146 TVKYTQGSWTGFGEDLVTIPKGFNTSFLVNIAT---IFESENFFLPGI-----KWNIGIL 197
      : |  || :|:: :| |::|  : |  |  ||  |  ||  |::|||
Db     139 DIHYGSGSLSGYLSQDTVSVPCSDQSKARGIKVEKQIF-GEATKQPGIVFVAAKFDGIL 197

Qy     198 GLAYATLAKPSSSLETFFDSLVTQANI-PNVFSMQMCGAGLPVAGSGTNGGSLVLGGIEP 256
      |: |  ::  ::  ||::| :  :  |::|  |  |  ||  |::||| :
Db     198 GMGYPHIS--VNNVLPVFDNLMQQKLVDKNIFSFY-----LNRDPEGQPGGELMLGGTDS 250

Qy     257 SLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVF 316
      | |:: |  :  : |::| :  :||:| :  |  |  |  :|||:|::||  | :
Db     251 KYYHGELSYLNVTRKAYWQVHMDQLEVGNE-LTL-CK--GGCEAIVDTGTSLLVGPVEEV 306

Qy     317 DAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRRSFRITIL 376
      : :|:  ||  :  |  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db     307 KELQKAIGAVPLI-----QGEYMIPCEKVSSL 333

Qy     377 PQLYIQPMMGAGLNYEC----YRFGIS-----PSTNALVIGATVMEG 414
      | :|:: :| | |||  |  :|  ||  :  :|  :
Db     334 PTVYLK--LG-GKNYELHPDKYILKVSQGGKTICLSGFMGMDIPPPSGPLWILGDVFIGS 390

Qy     415 FYVIFDRAQKRVGFA 429
      :| :|||  |||||
Db     391 YYTVFDRDNNRVGFA 405
  
```

Search completed: March 4, 2004, 15:40:59
 Job time : 29.1043 secs

OM protein - protein search, using sw model

Run on: March 4, 2004, 15:30:05 ; Search time 28.1043 Seconds
(without alignments)
1772.942 Million cell updates/sec

Title: US-09-668-314C-2
Perfect score: 2687
Sequence: 1 MGALARALLPLLAQWLLRA.....RPRDPEVVNDESSLVRHRWK 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_78:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	1178.5	43.9	501	2	A59090	aspartic proteinas
2	367.5	13.7	383	2	JC7573	pepsinogen C - Afr
3	363.5	13.5	377	1	PEMQCJ	gastricsin (EC 3.4
4	355.5	13.2	384	2	A39314	gastricsin (EC 3.4
5	355	13.2	389	2	JE0371	pepsin C (EC 3.4.2
6	353	13.1	388	2	A29937	gastricsin (EC 3.4
7	351.5	13.1	388	2	JC7246	pepsinogen C - com
8	324.5	12.1	394	2	B43356	gastricsin (EC 3.4
9	320	11.9	385	2	JC7575	pepsinogen A - bul
10	320	11.9	402	1	REMSK	renin (EC 3.4.23.1
11	313.5	11.7	509	2	S66516	oryzasin (EC 3.4.2
12	313	11.6	392	1	A24608	gastricsin (EC 3.4
13	310	11.5	383	2	A41443	pepsin (EC 3.4.23.

14	308.5	11.5	412	1	KHHUD	cathepsin D (EC 3.
15	306.5	11.4	410	1	KHMSD	cathepsin D (EC 3.
16	305.5	11.4	401	1	REMSS	renin (EC 3.4.23.1
17	305	11.4	384	2	JC7574	pepsinogen A - Afr
18	305	11.4	407	1	KHRTD	cathepsin D (EC 3.
19	302	11.2	405	2	A25379	saccharopepsin (EC
20	301.5	11.2	398	2	S66465	cathepsin E (EC 3.
21	300.5	11.2	387	2	C38302	pepsin (EC 3.4.23.
22	299	11.1	398	2	I51185	cathepsin D (EC 3.
23	298.5	11.1	387	2	D38302	pepsin (EC 3.4.23.
24	298.5	11.1	400	2	I47099	renin (EC 3.4.23.1
25	297	11.1	388	1	PEHU	pepsin A (EC 3.4.2
26	296	11.0	388	2	A30142	pepsin A (EC 3.4.2
27	296	11.0	388	2	B30142	pepsin A (EC 3.4.2
28	294.5	11.0	388	1	S19684	pepsin A (EC 3.4.2
29	292	10.9	506	2	T07915	probable aspartic
30	291	10.8	388	1	S19682	pepsin A (EC 3.4.2
31	291	10.8	402	1	RERTK	renin (EC 3.4.23.1
32	291	10.8	406	1	REHUK	renin (EC 3.4.23.1
33	290.5	10.8	396	2	S36865	cathepsin E (EC 3.
34	289	10.8	387	2	E38302	pepsin (EC 3.4.23.
35	288	10.7	387	2	B38302	pepsin (EC 3.4.23.
36	288	10.7	388	1	PEMQAJ	pepsin A (EC 3.4.2
37	287.5	10.7	632	2	T45858	hypothetical prote
38	287	10.7	391	2	A43356	cathepsin E (EC 3.
39	287	10.7	396	2	A34401	cathepsin E (EC 3.
40	286.5	10.7	334	2	JC4870	pepsin A (EC 3.4.2
41	286	10.6	382	1	PECH	pepsin A (EC 3.4.2
42	286	10.6	388	1	PEMQAR	pepsin A (EC 3.4.2
43	285.5	10.6	387	2	JC7245	pepsinogen A - com
44	285	10.6	396	2	T47207	aspartic proteinas
45	284.5	10.6	386	1	PEPG	pepsin A (EC 3.4.2

ALIGNMENTS

RESULT 1

A59090

aspartic proteinase (EC 3.4.23.-) BACE precursor - human

N;Alternate names: beta-secretase; beta-site APP cleaving enzyme

C;Species: Homo sapiens (man)

C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-May-2000

C;Accession: A59090

R;Vassar, R.; Bennett, B.D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E.A.; Denis, P.; Teplow, D.B.; Ross, S.; Amarante, P.; Loeloff, R.; Luo, Y.; Fisher, S.; Fuller, J.; Edenson, S.; Lile, J.; Jarosinski, M.A.; Biere, A.L.; Curran, E.; Burgess, T.; Louis, J.C.; Collins, F.; Treanor, J.; Rogers, G.; Citron, M.

Science 286, 735-741, 1999

A;Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.

A;Reference number: A59090; MUID:20002972; PMID:10531052

A;Note: submitted to GenBank, September 1999

A;Accession: A59090

A;Status: not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-501 <VAS>

A;Cross-references: GB:AF190725; NID:g6118538; PIDN:AAF04142.1; PID:g6118539
 C;Genetics:
 A;Gene: BACE
 C;Superfamily: beta-secretase
 C;Keywords: Alzheimer's disease; aspartic proteinase; brain; glycoprotein;
 hydrolase; protein digestion; transmembrane protein; zymogen
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-45/Domain: propeptide #status predicted <PRO>
 F;46-501/Product: acid proteinase BACE #status predicted <MAT>
 F;461-477/Domain: transmembrane #status predicted <TRN>
 F;93,289/Active site: Asp #status predicted
 F;153,172,223,354/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F;330-380/Disulfide bonds: #status predicted

Query Match 43.9%; Score 1178.5; DB 2; Length 501;
 Best Local Similarity 46.2%; Pred. No. 3.8e-80;
 Matches 240; Conservative 82; Mismatches 164; Indels 33; Gaps 9;

Qy	7	ALLLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLA	61
		: :	
Db	2	AQALPWLLLWM---GAGVLPAGHTQHGI RLPLRSGLG-----GAPL-----GLR	42
Qy	62	LALE--PALASPAGAAANFLAMVDNLQGDSSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVA	119
		: : : : : :	
Db	43	LPRETDEEPEEPGRGGSFVEMVDNLRGKSGQGYVEMTVGSPPTLNILVDTGSSNFAVG	102
Qy	120	GTPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIAT	179
		:: : : : : :	
Db	103	AAPHPFLHRY YQRQLSSTYRDLRKG VYVPYTQGWEGELGTDLVSI PHGPNVTVRANIAA	162
Qy	180	IFESENFFLPGIKWNGILGLAYATLAKPSSSLETF FDSLVTQANI PNVFSMQMCGAGLPV	239
		: : : : : : : :	
Db	163	ITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDSL VKQTHV PNLFSLQLCGAGFPL	222
Qy	240	AGS---GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWY YQIEILKLEIGGQSLNLD CREYN	296
		: :: : : : : :: : :	
Db	223	NQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWY YEVII VRVEINGQDLKMDCKEYN	282
Qy	297	ADKAIVDSGTTLLRLPQKVFD AVEAVARASLIPEFS DGFWTGSQLACWTNSETPWSYFP	356
		: : : :: : :	
Db	283	YDKSIVDSGTTNLR LPKKVFEAAVKS IKAASSTEKF PDGFWLGEQLVCWQAGTTPWNIFP	342
Qy	357	KIS IYLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGF	415
		: :: :: : : : : : :	
Db	343	VISLYLMGEVTNQSF RITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGF	402
Qy	416	YVIFDRAQKRVGFAASPCA EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYAL	475
		: : : : : : : :	
Db	403	YVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDESTLMTIAYVM	462
Qy	476	MSVCGAILLV LIVLLLLPFR CQR--RPRDPEVVNDESSL	512
		:: : :: : : : : : :	
Db	463	AAIC-ALFMLPLCLMVCQWRCLRCLRQQHDDFADDISLL	500

pepsinogen C - African clawed frog
 N;Alternate names: progastricsin
 C;Species: *Xenopus laevis* (African clawed frog)
 C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
 C;Accession: JC7573; PC7118
 R;Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
 J. Biochem. 129, 147-153, 2001
 A;Title: Amphibian pepsinogens: Purification and characterization of *Xenopus* pepsinogens, and molecular cloning of *Xenopus* and bullfrog pepsinogens.
 A;Reference number: JC7573; MUID:21064922; PMID:11134969
 A;Contents: Stomach
 A;Accession: JC7573
 A;Molecule type: mRNA
 A;Residues: 1-383 <IKU>
 A;Cross-references: DDBJ:AB045379
 A;Accession: PC7118
 A;Molecule type: protein
 A;Residues: 17-68 <IK2>
 C;Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like activity.
 C;Genetics:
 A;Gene: PgC
 C;Superfamily: pepsin
 C;Keywords: stomach; zymogen

Query Match 13.7%; Score 367.5; DB 2; Length 383;
 Best Local Similarity 28.9%; Pred. No. 9.1e-20;
 Matches 132; Conservative 70; Mismatches 154; Indels 101; Gaps 25;

Qy	1	MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
		: :: :: : : :::	
Db	1	MKFLILALVCLQLSEGIIR-----VPLKKFKSMREVMRENGIKAPLVDPAT---KYYNQY	52
Qy	61	ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
		: :: : :	
Db	53	ATAYEP-----LSNYMDM-----SYYGEISIGTPPQNFLVLFDTGSSNLWVAS	95
Qy	121	TPHSYIDT-----YFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSF	173
		: : :: :	
Db	96	T---YCQSQACTNHPLFNPSQSSTYSSNQQQFSLQYGTGSLTGILGYDTVTIQ-----	145
Qy	174	LVNIATIFESENFEL----PG-----IKWNGILGLAYATLAKPSSSLETFFDSLVTQANI	224
		: : :: : :: :: :	
Db	146	--NVA--ISQQEFGLETEPGTNFVYAQFDGILGLAYPSIAVGGAT--TVMQGM--QQNL	198
Qy	225	PN--VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLE	282
		: : : :: : :: : :	
Db	199	LNQPIFGFYLSGQ-----SSQNGGEVAFGGVDQNYTGTQIYWTPVTSETYWQIGIQGFS	252
Qy	283	IGGQSLNLDREYNADKAIVDSGTTLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ	342
		: : : : : ::::: : :	
Db	253	INGQATGW-CSQ--GCQAIVDTGTSLLTAPQSVFSSLIQSIG-----AQQDQNGQYV	301
Qy	343	ACWTNSETPWSYFPKISIIYLRDENSSRSFRITILPQLYI-QPMMGAGLNYECYRFGIS--	399
		: : : : :	
Db	302	VSCSNIQN----LPTISFTI----SGVSFPLP--PSAYVLQQSSG-----YC-TIGIMPT	345

Qy 400 --PSTNA---LVIGATVMEGFYVIFDRAQKRVGFAAS 431
 || | :| | : :| :| | :| | | | :
 Db 346 YLPSQNGQPLWILGDVFLREYYSVYDLGNNQVGFATA 382

RESULT 3

PEMQCJ

gastricsin (EC 3.4.23.3) precursor - Japanese macaque (fragment)

N;Alternate names: pepsin C

C;Species: *Macaca fuscata* (Japanese macaque)

C;Date: 13-Aug-1986 #sequence_revision 19-Oct-1995 #text_change 18-Jun-1999

C;Accession: S19683; A00986; A22402; S16066

R;Kageyama, T.; Tanabe, K.; Koiwai, O.

Eur. J. Biochem. 202, 205-215, 1991

A;Title: Development-dependent expression of isozymogens of monkey pepsinogens and structural differences between them.

A;Reference number: S19681; MUID:92037645; PMID:1935977

A;Accession: S19683

A;Molecule type: mRNA

A;Residues: 1-377 <KAG>

A;Cross-references: EMBL:X59754; NID:g38072; PIDN:CAA42426.1; PID:g38073

R;Kageyama, T.; Takahashi, K.

J. Biol. Chem. 261, 4406-4419, 1986

A;Title: The complete amino acid sequence of monkey progastricsin.

A;Reference number: A00986; MUID:86168133; PMID:3514597

A;Accession: A00986

A;Molecule type: protein

A;Residues: 6-330, 'V', 332-349, 'VY', 350-377 <KA2>

R;Kageyama, T.; Takahashi, K.

J. Biochem. 97, 1235-1246, 1985

A;Title: Monkey pepsinogens and pepsins. VII. Analysis of the activation process and determination of the NH2-terminal 60-residue sequence of Japanese monkey progastricsin, and molecular evolution of pepsinogens.

A;Reference number: A22402; MUID:85289106; PMID:3928607

A;Accession: A22402

A;Molecule type: protein

A;Residues: 6-65 <KA3>

C;Comment: This enzyme has more restricted specificity than pepsin A.

C;Comment: The enzyme is activated in a two-step process that gives rise to two end products. The shorter, Ser-gastricsin, is the major product.

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach

F;1-5/Domain: signal sequence (fragment) #status predicted <SIG>

F;6-377/Product: progastricsin #status experimental <ZYM>

F;6-45/Domain: activation peptide #status experimental <APT>

F;46-377/Product: Gly-gastricsin #status experimental <MIN>

F;49-377/Product: Ser-gastricsin #status experimental <MAT>

F;31-32/Cleavage site: Phe-Leu (pepsin) #status experimental

F;45-46/Cleavage site: Phe-Gly (pepsin) #status experimental

F;48-49/Cleavage site: Leu-Ser (pepsin) #status experimental

F;80,265/Active site: Asp #status predicted

F;93-98,256-260,299-332/Disulfide bonds: #status experimental

Query Match 13.5%; Score 363.5; DB 1; Length 377;

Best Local Similarity 28.9%; Pred. No. 1.8e-19;

Matches 118; Conservative 65; Mismatches 118; Indels 107; Gaps 19;

Qy 56 HADGLALALEPALASPAGANFLAMVDNLQDGRGYYLEMLIGTPPQKLQILVDTGSSN 115
| |::: || :| :| | : | : ||||| :| |||||
Db 44 HFGDLSVSYEP-----MAYMD-----AAYFGEISIGTPPQNFLVLFDTGSSN 85

Qy 116 FAV-----AGTPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGEVGEDLVITPK 167
| | | | | : | : : | | | | | : | :
Db 86 LWVPSVYCQSQACTSHS----RFPSESTYSTNGQTFSLQYGSGLTGFFGYDTLTV-- 139

Qy 168 GFNTSFLVNIATIFESENFPLPG-----IKWNGILGLAYATLAKPSSSLETFFDSLVTQA 222
| | | | | : : : | : | : | : | : | :
Db 140 ---QSIQVPNQEFGLSEN--EPGTNEFVYAQFDGIMGLAYPTLSVDGAT--TAMQGMVQEG 192

Qy 223 NIPN-VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKL 281
: : : | : : | : : | : | : | : | : | :
Db 193 ALTSPIFSVELSDQ-----QGSSGAVVFGVDSSLYTGQIYWAPVTQELYWQIGIEEF 246

Qy 282 EIGGQSLNLDREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ 341
| | | : | | : | | : | : | : | : | :
Db 247 LIGGQASGW-CSE--GCQAIVDTGTSLTVPQQYMSALLQA-----TGAQ 288

Qy 342 LACWTNSETPWSYF-----PKISYLRDENSSRSFRITILPQLYIQPMGAGLNY 391
| : | | : : : | : | : | : | :
Db 289 -----EDEYGOFLVNCNSIQNLPTLTFTII-----NGVEFPLPPSSYI-----LNN 328

Qy 392 EGY-RFGISP-----STNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
| | : | | : : | : : | : | : | :
Db 329 NGYCTVGVEPTYLSAQNSQPLWILGDVFLRSYYSVYDLSNNRVGFATA 376

RESULT 4

A39314

gastricsin (EC 3.4.23.3) precursor - bullfrog

C;Species: Rana catesbeiana (bullfrog)

C;Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 22-Jun-1999

C;Accession: A39314

R;Yakabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Miki, K.; Kurokawa, K.; Ito, H.; Kageyama, T.; Takahashi, K.

J. Biol. Chem. 266, 22436-22443, 1991

A;Title: Purification, characterization, and amino acid sequences of pepsinogens and pepsins from the esophageal mucosa of bullfrog (Rana catesbeiana).

A;Reference number: A39314; MUID:92042186; PMID:1939266

A;Accession: A39314

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-384 <YAK>

A;Cross-references: GB:M73750; NID:g213687; PIDN:AAA49530.1; PID:g213688

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; hydrolase; protein digestion

Query Match 13.2%; Score 355.5; DB 2; Length 384;
Best Local Similarity 26.5%; Pred. No. 7.2e-19;
Matches 120; Conservative 73; Mismatches 136; Indels 123; Gaps 21;

Qy 23 ELAPAPFTLPLRVAAATNRVV-----APTPGPGTPAERHADGLALALEPALASPAGAN 76
: | : : | : : | : : : | | | | |
Db 12 QLSEGIKIVPLKKFKSMREVMRDHGIKAPVVDPAT---KYNNFATAFEP-----LAN 61

```

Qy      77 FLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDT-----Y 129
      :: |          || |: ||||| :| ||||| | | :| :
Db      62 YMDM-----SYIGEISIGTPPQNFLVLEDTGSSNLWV---PSTYCQSQACTNHPQ 108

Qy     130 FDTERSSTYRSKGFDTVKYTGQSWTGFVGEDLV TIPKGFNTSFLVNIATIFESENFLL- 188
      |: :||:| | :::| || || :| | | | ||| : | |
Db     109 FNPSQSSSYSSNQQQFSLQYGTGSLTGILGYDTVQIQ-----NIA--ISQQEFGLS 157

Qy     189 ---PG-----IKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPN--VFMSQMCGAGLP 238
      || :::||||| :||: :: | : : | | : | :|: : |
Db     158 VTEPGTNFVYAQFDGILGLAYPSIAEGGAT--TVMQGMII-QQNLIHQPLFAFYLSG---- 210

Qy     239 VAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNAD 298
      : ||| : ||:: : | | |::||: | |:|| | : ||: | :
Db     211 -QQNSQNGGEVAFGGVDQNYYSGQIYWTPVTSETYWQIGIQGFSVNGQATGW-CSQ--GC 266

Qy     299 KAIVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ-----LACWTNSET 350
      : |||:||:| | || || ::::: | :| : :|
Db     267 QGIVDTGTSLLTAPQSVFSSLMQSI-----GAQQDQNGQYAVSCSNIQSL 311

Qy     351 PWSYFP-----KISIIYLRDENS---SRSFRTILPQLYIQPMMGAGLNYECYREGIS 399
      | | | | : | : || : | | | ||:
Db     312 PTISFTISGVSFPLPPSAYVLQQNSGYCTIGIMPTYLPSQNGQPLW----- 357

Qy     400 PSTNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
      ::| : :| ::| :|||:
Db     358 -----ILGDVFLRQYYSVYDLGNNQVGFAAA 383

```

RESULT 5

JE0371

pepsin C (EC 3.4.23.-) precursor - chicken

N;Alternate names: pepsinogen C

C;Species: Gallus gallus (chicken)

C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000

C;Accession: JE0371

R;Sakamoto, N.; Saiga, H.; Yasugi, S.

Biochem. Biophys. Res. Commun. 250, 420-424, 1998

A;Title: Analysis of temporal expression pattern and cis-regulatory sequences of chicken pepsinogen A and C.

A;Reference number: JE0370; MUID:98440813; PMID:9753645

A;Accession: JE0371

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-389 <SAK>

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; hydrolase

Query Match 13.2%; Score 355; DB 2; Length 389;

Best Local Similarity 28.7%; Pred. No. 7.9e-19;

Matches 114; Conservative 58; Mismatches 121; Indels 104; Gaps 16;

```

Qy      75 ANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGT-----PHSYI 126
      :|| : | : || |: ||||| :| ||||| | | | :
Db      56 SNFATAYEPLANNDMSYYGEISIGTPPQNFLVLEDTGSSNLWVPSTLCQSQACANHN-- 113

```

Qy 127 DTYFDTERSSTYRSKGFDTVKYTGGSWTGFVGEDLVTIPKGFN-----TS 172
 || |||: :: ::| || || | ||| :| : ||
 Db 114 --EFDPNESSTFSTQDEFFSLQYGSGLTGIFGFDTVTI-QGISITNQEFGLSETEPGTS 170

Qy 173 FLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPN--VFSM 230
 || : ::||||||: :: |: | :| : |||
 Db 171 FLYS-----PFDGILGLAFPSI---SAGGATTVMQKMLQENLLDFPVFSF 212

Qy 231 QMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNL 290
 : | |: || || ||::|:| | | :||: : |:| | :|||
 Db 213 YLSGQ-----EGSQGGELVFGGVDPNLYTGQITWTPVTQTTYWQIGIEDFAVGGQSSGW 266

Qy 291 DCREYNADKAIVDSGTTLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSET 350
 | : : |||:|:| | :| :| : :: : : | :| :| :| |
 Db 267 -CSQ--GCQGIVDGTGTSLLTVPNQVFTELMQYIG-----AQADD---SGQYVASCNIE- 314

Qy 351 PWSYFPKI-----SIYLRDENS---SRSFRITILPQLYIQPMMGAGLNYECY 394
 | | | | | : || : | || ||:
 Db 315 ---YMPTITFVISGTSFPLPPSAYMLQSNSDYCTVGIESTYLPSQTGQPLW----- 362

Qy 395 RFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
 :| : :| | :| :||| :
 Db 363 -----ILGDVFLRVYYSIYDMGNNQVGFATA 388

RESULT 6

A29937

gastricsin (EC 3.4.23.3) precursor - human

N;Alternate names: pepsin C; pepsinogen C

C;Species: Homo sapiens (man)

C;Date: 17-Oct-1988 #sequence_revision 17-Oct-1988 #text_change 31-Mar-2000

C;Accession: A29937; A31811; PX0028; I54213; A91125; A23458

R;Hayano, T.; Sogawa, K.; Ichihara, Y.; Fujii-Kuriyama, Y.; Takahashi, K.

J. Biol. Chem. 263, 1382-1385, 1988

A;Title: Primary structure of human pepsinogen C gene.

A;Reference number: A29937; MUID:88087276; PMID:3335549

A;Accession: A29937

A;Molecule type: DNA

A;Residues: 1-388 <HAY>

R;Taggart, R.T.; Cass, L.G.; Mohandas, T.K.; Derby, P.; Barr, P.J.; Pals, G.; Bell, G.I.

J. Biol. Chem. 264, 375-379, 1989

A;Title: Human pepsinogen C (progastricsin). Isolation of cDNA clones, localization to chromosome 6, and sequence homology with pepsinogen A.

A;Reference number: A31811; MUID:89079679; PMID:2909526

A;Accession: A31811

A;Molecule type: mRNA

A;Residues: 1-388 <TAG>

A;Cross-references: GB:J04443; NID:g551175; PIDN:AAA60074.1; PID:g551176

R;Athauda, S.B.P.; Tanji, M.; Kageyama, T.; Takahashi, K.

J. Biochem. 106, 920-927, 1989

A;Title: A comparative study on the NH2-terminal amino acid sequences and some other properties of six isozymic forms of human pepsinogens and pepsins.

A;Reference number: PX0023; MUID:90130402; PMID:2515193

A;Accession: PX0028

A;Molecule type: protein

A;Residues: 17-101 <ATH>

R;Pals, G.; Azuma, T.; Mohandas, T.K.; Bell, G.I.; Bacon, J.; Samloff, I.M.; Walz, D.A.; Barr, P.J.; Taggart, R.T.
 Genomics 4, 137-148, 1989
 A;Title: Human pepsinogen C (progastricsin) polymorphism: evidence for a single locus located at 6p21.1-pter.
 A;Reference number: I54213; MUID:89290840; PMID:2567697
 A;Accession: I54213
 A;Status: translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-388 <RES>
 A;Cross-references: GB:M23077; NID:g189830; PIDN:AAA60063.1; PID:g387015; GB:J03063
 A;Note: parts of this sequence, including the amino end and carboxyl ends of the mature protein, were determined by protein sequencing
 R;Foltmann, B.; Jensen, A.L.
 Eur. J. Biochem. 128, 63-70, 1982
 A;Title: Human progastricsin. Analysis of intermediates during activation into gastricsin and determination of the amino acid sequence of the propeptide.
 A;Reference number: A91125; MUID:83079318; PMID:6816595
 A;Accession: A91125
 A;Molecule type: protein
 A;Residues: 17-39, 'ED', 42-51, 'S', 53-64 <FOL>
 A;Note: pro-form; 29-Leu was also found
 A;Note: activation at pH 2 is proposed to involve conformation change, cleavage after Phe-42, and cleavage after Leu-59
 C;Genetics:
 A;Gene: GDB:PGC
 A;Cross-references: GDB:119485; OMIM:169740
 A;Map position: 6p21.3-6p21.1
 A;Introns: 20/2; 70/3; 110/1; 149/3; 216/2; 256/2; 305/3; 338/3
 C;Superfamily: pepsin
 C;Keywords: aspartic proteinase; hydrolase; protein digestion; stomach; zymogen
 F;1-16/Domain: signal sequence #status predicted <SIG>
 F;17-59/Domain: propeptide #status experimental <PRO>
 F;60-388/Product: gastricsin #status experimental <MAT>

Query Match 13.1%; Score 353; DB 2; Length 388;
 Best Local Similarity 29.1%; Pred. No. 1.1e-18;
 Matches 120; Conservative 65; Mismatches 120; Indels 108; Gaps 21;

Qy	52	PAERHADG-LALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVD	110
		:: :: : : : : :	
Db	50	PAWKYRFGDLSVTYEP-----MAYMD-----AAYFGEISIGTPPQNFLVLFD	91
Qy	111	TGSSNFAV-----AGTPHSYIDTYFDTERSSTYRSKGFDTVYKYTGGSWTGFVGEDL	162
		: : ::	
Db	92	TGSSNLWVPSVYCQSQACTSHS----RFNPSESSTYSTNGQTFSLQYGSGLTGFFGYDT	147
Qy	163	VTIPKGFNTSFLVNIATIFESENFLLPG-----IKWNGILGLAYATLAKPSSSLETFFDS	217
		: : : : : : :	
Db	148	LTV-----QSIQVPNQEFGLSEN--EPGTNFVYAQFDGIMGLAYPALSVDEAT--TAMQG	198
Qy	218	LVTQANIPN-VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQI	276
		: : : : : : : : : : :	
Db	199	MVQEGALTSPVFSVYLSNQ-----QGSSGGAUVFGVDSSLYTGQIYWAPVTQELYWQI	252
Qy	277	EILKLEIGGQSLNLDREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGF	336

```

      | : |||: | | :|||:|:| | :||: |::|
Db      253 GIEEFLIGGQASGW-CSE--GCQAIVDTGTSLLTVPQQYMSALLQA----- 295

Qy      337 WTGSQLACWTNSETPWSYF-----PKISIIYLRDENSRRSFRITILPQLYIQPMMG 386
      ||:| | : | | : : | : | |
Db      296 -TGAQ-----EDEYGQFLVNCNSIQNLPSLTFII-----NGVEFPLPPSSYI----- 336

Qy      387 AGLNYECY-RFGISP----STNA---LVIGATVMEGFYVIFDRAQKRVGFAAS 431
      |: | |: | | :| : :| :| :| | ||| :
Db      337 --LSNNGYCTVGVEPTYLSSQNGQPLWILGDVFLRSYYSVYDLGNNRVGFATA 387

```

RESULT 7

JC7246

pepsinogen C - common marmoset

C;Species: Callithrix jacchus (common marmoset)

C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000

C;Accession: JC7246

R;Kageyama, T.

J. Biochem. 127, 761-770, 2000

A;Title: New world monkey pepsinogens A and C, and prochymosins. Purification, characterization of enzymatic properties, cDNA cloning, and molecular evolution.

A;Reference number: JC7245

A;Accession: JC7246

A;Molecule type: mRNA

A;Residues: 1-388 <KAG>

A;Cross-references: DDBJ:AB038385

A;Experimental source: strain NW791

C;Comment: This protein, a zymogen of pepsins, is the major proteolytic enzyme in vertebrate gastric juices. It plays roles in gastric digestion, and is a useful molecular marker for clarifying the evolution of mammalian orders and families.

C;Superfamily: pepsin

C;Keywords: gastric juice; zymogen

```

Query Match          13.1%;  Score 351.5;  DB 2;  Length 388;
Best Local Similarity 30.1%;  Pred. No. 1.4e-18;
Matches 112;  Conservative 56;  Mismatches 115;  Indels 89;  Gaps 17;

```

```

Qy      92 YYLEMLIGTPPQKLQILVDTGSSNFAV-----AGTPHSYIDTYFDTERSSTYRSKGF 143
      |: |: ||||| :| ||||| | | | | |: | ||| | |
Db      73 YFGEISIGTPPQNFLVLFDTGSSNLWVPSVYCQSQACTSHS----RFNPSASSTYSSNGQ 128

Qy      144 DVTVKYTQGSWTGFVGEDLVTI PKGENTSFLVNIATIFESENFFLPG----IKWNGILG 198
      :::| || ||| | | :|: | | | | || | | :::||:|
Db      129 TFSLQYGSGLTGFFGYDTLTV-----QSIQVPNQEFGLSEN--EPGTNFVYAQFDGIMG 181

Qy      199 LAYATLAKPSSSLETFFDSLVTQANIPN-VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPS 257
      ||| |: : : | : : : : ||| : |::||::: ||:: |
Db      182 LAYPALSMGGAT--TAMQGM LQEGALTSPVFSFYLSNQ-----QGSSGGAVIFGGVDSS 233

Qy      258 LYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKAIVDSGTLLRLPQKVFD 317
      || | |:: |: :| |::| | : |||: | | :|||:|:| | :||:
Db      234 LYTGQIYWAPVTQELYWQIGIEEFLIGGQASGW-CSE--GCQAIVDTGTSLLTVPQQYMS 290

Qy      318 AVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYF-----PKISIIYLRDENS 367
      | :|| |::| | : | | : :

```

```

Db      291 AFLEA-----TGAQ-----EDEYGQFLVNCDSIQNLPTLTFII----- 323
Qy      368 SRSFRITILPQLYIQPMMGAGLNYECY-RFGISP-----STNALVIGATVMEGFYVIF 419
          : | ||      | : | | : |      | : | : | : |
Db      324 -NGVEFPLPPSSYI-----LSNNGYCTVGVEPTYLSSQNSQPLWILGDVFLRSYYSVF 375
Qy      420 DRAQKRVGFAAS 431
          | |||| :
Db      376 DLGNNRVGFATA 387

```

RESULT 8

B43356

gastricsin (EC 3.4.23.3) precursor - guinea pig

N;Alternate names: pepsin C

C;Species: Cavia porcellus (guinea pig)

C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999

C;Accession: B43356

R;Kageyama, T.; Ichinose, M.; Tsukada, S.; Miki, K.; Kurokawa, K.; Koiwai, O.;
TANJI, M.; Yakabe, E.; Athauda, S.B.; Takahashi, K.
J. Biol. Chem. 267, 16450-16459, 1992

A;Title: Gastric procathepsin E and progastricsin from guinea pig. Purification,
molecular cloning of cDNAs, and characterization of enzymatic properties, with
special reference to procathepsin E.

A;Reference number: A43356; MUID:92355614; PMID:1644829

A;Accession: B43356

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-394 <KAG>

A;Cross-references: GB:M88652; NID:g191296; PIDN:AAA37053.1; PID:g191297

A;Note: sequence extracted from NCBI backbone (NCBIN:110805, NCBIP:110806)

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion;
stomach

```

Query Match          12.1%;  Score 324.5;  DB 2;  Length 394;
Best Local Similarity 29.0%;  Pred. No. 1.5e-16;
Matches 107;  Conservative 63;  Mismatches 116;  Indels 83;  Gaps 18;

```

```

Qy      92 YYLEMLIGTPPQKLQILVDTGSSNF-----AVAGTPHSYIDTYFDTERSSTYRSKGF 143
          |: :: :|||| | :| |||||      ::| | | | | : | ||| :
Db      79 YFGQISLGTPPQSFQVLFDTGSSNLWVPSVYCSSLACTTH----TRFNPRDSSTYVATDQ 134
Qy      144 DVTVKYTQGSWTGFVGEDLVTI-----PK-GFNTSFLVNIATIFESENFFLPG-----IK 192
          :::| || || | | :||      || | | | | :| | | :
Db      135 SFSLEYGTGSLTGVFGYDTMTIQDIQVPKQEFGLS-----ETE----PGSDFVYAE 181
Qy      193 WNGILGLAYATLAKPSSSLETFFDLVTQANI-PNVFSMQMCGAGLPVAGS--GTNGGSL 249
          ::|||| | | : : : | | : : : : ||| : : | | : |
Db      182 FDGILGLGYPLSEGGAT--TAMQGLLREGALSQSLFSVYL-----GSQQGSDEGQL 231
Qy      250 VLGgiePSLYKGDWYTPiKEEWYYQIEILKLEIGGQSLNLDcreYNADKAIVDSGTTLL 309
          :|||: : ||| |||: ||: :| | :| | | | : | : |||: ||: ||
Db      232 ILGGVDESlyTgDIYwTPtQELyWQIGIEGFLIDGSASGWCSR---GCQGIVDTGTSLL 288
Qy      310 RLPQKVFDaVVEaVaRaSLiPEfSDGFWtGSQLaCWtNSeTPwSYfPKiSiYLRdENSSR 369
          :|      :|:| : | : : | : : | : : | :

```

```

Db      289 TVPSDYLSTLVQAIGAE--NEYGEYF-----VSCSSIQDLPTLTFVISGV----- 332
Qy      370 SFRITILPQLYIQP-----MMGAGLNYECYRFGISPSTN--ALVIGATVMEGFYVIFDRA 422
          : | ||      | : |      |      : ||      : : |      : : | : : | |
Db      333 --EFPLSPSAYILSGENYCMVGLESTY-----VSPGGGEPVWILGDVFLRSYYSVYDLA 384
Qy      423 QKRVGFAAS 431
          ||||| :
Db      385 NNRVGFATA 393

```

RESULT 9

JC7575

pepsinogen A - bullfrog

C;Species: Rana catesbeiana (bullfrog)

C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001

C;Accession: JC7575

R;Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.

J. Biochem. 129, 147-153, 2001

A;Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinogens, and molecular cloning of Xenopus and bullfrog pepsinogens.

A;Reference number: JC7573; MUID:21064922; PMID:11134969

A;Contents: Stomach

A;Accession: JC7575

A;Molecule type: mRNA

A;Residues: 1-385 <IKU>

A;Cross-references: DDBJ:AB045376

C;Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-like activity.

C;Genetics:

A;Gene: PgA

C;Superfamily: pepsin

C;Keywords: stomach; zymogen

```

Query Match          11.9%;  Score 320;  DB 2;  Length 385;
Best Local Similarity 27.8%;  Pred. No. 3.2e-16;
Matches 111;  Conservative 67;  Mismatches 147;  Indels 74;  Gaps 15;

```

```

Qy      50 GTPAERHADGLALALEPALASPAGAAFLAMVDNLQDSEGRGYYLEMLIGTPPQKLQILV 109
          | : : |      |      | : ||      : |      : ||      | : : ||||| : :
Db      39 GDYLKKHHYNPATKYFPSLAQASG-----EPLQNYMDIEYFGTISIGTPPQSFTVIF 90
Qy      110 DTGSSNFAVAGTPHSYIDT-----YFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDL 162
          ||||| |      | :      | : : ||| : :      | : : | || : || : |
Db      91 DTGSSNLWV---PSVYCSPACTNHHMFNPQSSSTFQATNTPVSIQYGTGMSGFLGYDT 147
Qy      163 VTIPKGFNTSFLVNIATIFESE--NFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQ 221
          | :      |      | : : ||      ||      : : ||||| : : ||      || : : |
Db      148 VQVG---NIQITNQIFGLSQSEPGSFLYYSFPDGILGLAFPSLA--SSQATPVFDNMWNQ 202
Qy      222 ANIP-NVFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILK 280
          || : : || : :      : | : | : : || : : | | | : : | : | : || :
Db      203 GLIPQDLFSVYL-----SSQGQSGSFVLFGGVDTSYTGNLNVVPLTAETYWQITVDS 255
Qy      281 LEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGS 340
          : |||| :      : |||| : || : ||      |      | : :      | :      : | : | :
Db      256 ISIGGQVIACS----GSCSAIVDTGTSLLAGPSTPI-ANIQYYIGAN---QDSNGQYV-- 305

```

```

Qy      341 QLACWTNSETPWSYFP-----KISILYLRDENSS--RSFRITILPQLYIQPMMGAGLN 390
          : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      306 -INCNNISNMPTVVFTINGVQYPLPASAYVRQSQSQSCTSGFQAMNLP----- 351

Qy      391 YECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFA 429
          : | : : | : : | | | | | | | | | | | | | | | | | | | | | |
Db      352 -----TSSGDLWILGDVFIREFYVVFDRANNYVAMA 382

```

RESULT 10

REMSK

renin (EC 3.4.23.15) precursor, renal - mouse

N;Alternate names: angiotensin-forming enzyme; angiotensinogenase; renin 1

C;Species: Mus musculus (house mouse)

C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 18-Jun-1999

C;Accession: A00989; S07636; A22766; A22058; I57576; A05137; JH0083

R;Holm, I.; Ollo, R.; Panthier, J.J.; Rougeon, F.

EMBO J. 3, 557-562, 1984

A;Title: Evolution of aspartyl proteases by gene duplication: the mouse renin gene is organized in two homologous clusters of four exons.

A;Reference number: A00989; MUID:84182525; PMID:6370686

A;Accession: A00989

A;Molecule type: DNA

A;Residues: 1-402 <HOL>

A;Cross-references: EMBL:X00850

R;Kim, W.S.; Murakami, K.; Nakayama, K.

Nucleic Acids Res. 17, 9480, 1989

A;Title: Nucleotide sequence of a cDNA coding for mouse Ren1 preprorenin.

A;Reference number: S07636; MUID:90067953; PMID:2685761

A;Accession: S07636

A;Molecule type: mRNA

A;Residues: 1-402 <KIM>

A;Cross-references: EMBL:X16642; NID:g53930; PIDN:CAA34636.1; PID:g53931

R;Mullins, J.J.; Burt, D.W.; Windass, J.D.; McTurk, P.; George, H.; Brammar, W.J.

EMBO J. 1, 1461-1466, 1982

A;Title: Molecular cloning of two distinct renin genes from the DBA/2 mouse.

A;Reference number: A90968; MUID:84207899; PMID:6327270

A;Accession: A22766

A;Molecule type: mRNA

A;Residues: 269-314, 'D', 316 <MUL>

R;Panthier, J.J.; Dreyfus, M.; Roux, D.T.L.; Rougeon, F.

Proc. Natl. Acad. Sci. U.S.A. 81, 5489-5493, 1984

A;Title: Mouse kidney and submaxillary gland renin genes differ in their 5' putative regulatory sequences.

A;Reference number: A22058; MUID:84298161; PMID:6089205

A;Accession: A22058

A;Molecule type: DNA

A;Residues: 1-30 <PAN>

R;Field, L.J.; Philbrick, W.M.; Howles, P.N.; Dickinson, D.P.; McGowan, R.A.; Gross, K.W.

Mol. Cell. Biol. 4, 2321-2331, 1984

A;Title: Expression of tissue-specific Ren-1 and Ren-2 genes of mice: Comparative analysis of 5'-proximal flanking regions.

A;Reference number: I57576; MUID:85085936; PMID:6392850

A;Accession: I57576

A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-31 <RES>
A;Cross-references: GB:K02800; NID:g200689; PIDN:AAA40044.1; PID:g200690
C;Comment: The only known function of renal renin is to release angiotensin I from angiotensinogen in the plasma, initiating a cascade of reactions that produces an elevation of blood pressure and increased sodium retention by the kidney.
C;Comment: Renal renin is synthesized by the juxtaglomerular cells of the kidney in response to decreased blood pressure and sodium concentration.
C;Genetics:
A;Gene: Ren-1
A;Introns: 31/2; 81/3; 123/1; 162/3; 228/2; 268/2; 316/3; 349/3
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; blood pressure control; glycoprotein; hydrolase; kidney; plasma
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-64/Domain: propeptide #status predicted <PRO>
F;65-402/Product: renin #status predicted <MAT>
F;69,139,320/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;102,287/Active site: Asp #status predicted

Query Match 11.9%; Score 320; DB 1; Length 402;
Best Local Similarity 28.6%; Pred. No. 3.4e-16;
Matches 126; Conservative 66; Mismatches 181; Indels 68; Gaps 21;

Qy	10	LPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPG-PGTPAERHADGLALALE---	65
		: : : : :	
Db	6	MPLWALLLL-----WSPCTFSLPTRTATFERIPLKKMPSVREILEERGVDMTRLSEAEGV	60
Qy	66	----PA---LASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAV	118
		: : : :: :	
Db	61	FTKRPSLTNLTSPVVLTYNL----NTQ-----YYGEIGIGTPPQTFKVIFDTGSANLWV	110
Qy	119	AGTPHSY-----IDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTS	172
		: :: : : : : : :	
Db	111	PSTKCSRLYLACGIHSLYESSDSSSYMENGSDFTIHYGSGRVKGFLSQDSVTV-GGITVT	169
Qy	173	FLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANI-PNVFSMQ	231
		: : : : :: : :	
Db	170	QTFGEVTELPLIPFML--AKFDGVLGMGFP--AQAVGGVTPVFDHILSQGVLKEEVFSVY	225
Qy	232	MCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD	291
		: : : : : : : : :	
Db	226	Y-----NRGSHLLGGEVVLGGSDPQHYQGNFHYVSISKTDWQITMKGVSVG--SSTLL	277
Qy	292	CREYNADKAIVDSGTTLLRLPQKVFDVAVVEAV-ARASLIPEFSDGFWTGSQACWTNSET	350
		: : : : : : : : :	
Db	278	CEEGCA--VVVDTGSSFISAPTSSLKLIMQALGAKEKRIEY-----VVNC---SQV	324
Qy	351	PWSYFPKISIIYLRDENSSRSFRITILPQLYIQPMMGAGL-NYECYRFGISPSTNAL-VIG	408
		: : : : : : :	
Db	325	P--TLPDISFDL----GGRAYTLSSTDYVLQYPNRRDKLCTLALHAMDIPPTGPVWVLG	378
Qy	409	ATVMEGFYVIFDRAQKRVGFA	429
		: :	
Db	379	ATFIRKFYTEFDRHNNRIGFA	399

RESULT 11

S66516

oryzasin (EC 3.4.23.-) precursor - rice

N;Alternate names: aspartic proteinase 1

C;Species: Oryza sativa (rice)

C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000

C;Accession: S66516; S66517

R;Asakura, T.; Watanabe, H.; Abe, K.; Arai, S.

Eur. J. Biochem. 232, 77-83, 1995

A;Title: Rice aspartic proteinase, oryzasin, expressed during seed ripening and germination, has a gene organization distinct from those of animal and microbial aspartic proteinases.

A;Reference number: S66516; MUID:96048031; PMID:7556174

A;Accession: S66516

A;Molecule type: DNA

A;Residues: 1-509 <ASA>

A;Cross-references: EMBL:D32165; NID:g511665; PIDN:BAA06876.1; PID:g1030715

A;Accession: S66517

A;Molecule type: mRNA

A;Residues: 1-509 <ASZ>

A;Cross-references: EMBL:D32144; NID:g1255684; PIDN:BAA06875.1; PID:g1711289

C;Comment: The pair of saposin repeat homology domains tagged SAP1 and SAP2 represent a cyclical permutation of a single saposin repeat.

C;Genetics:

A;Introns: 119/3; 140/1; 171/3; 209/2; 265/3; 279/1; 300/3; 338/3; 360/2; 412/3; 452/3; 482/2

C;Superfamily: oryzasin; saposin repeat homology

C;Keywords: aspartic proteinase; hydrolase

F;1-20/Domain: signal sequence #status predicted <SIG>

F;21-68/Domain: propeptide #status predicted <PRO>

F;68-509/Product: aspartic proteinase 1 #status predicted <MAT>

F;316-361/Domain: saposin repeat homology #status atypical <SAP1>

F;370-420/Domain: saposin repeat homology #status atypical <SAP2>

F;103,290/Active site: Asp #status predicted

Query Match 11.7%; Score 313.5; DB 2; Length 509;

Best Local Similarity 23.0%; Pred. No. 1.5e-15;

Matches 127; Conservative 75; Mismatches 179; Indels 171; Gaps 19;

```

Qy      3 ALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGLAL 62
      ::|  ||  :|  ||  |:  |      |  |      :||  |  |      |  |  |
Db      5 SVALVLLAAVLLQALLPASAEGLVRIALKKRPIDENSRVAARLSG-----EEGARRLGL 59

Qy     63 ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSN----- 115
      :|  |  :  :|:  :  :      |:  |:  :|||  |  :  :|||  |
Db     60 RGANSLGGGGGEGDIVALKNYMNAQ----YFGEIGVGTPPQKFTVIFDTGSSNLWVPSAK 115

Qy    116 --FAVAGTPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNTSF 173
      |::|  ||      :  :  :|||  :  |      ::|  ||  ||  ||  ||:
Db    116 CYFSIACFFHS----RYKSGQSSTYQKNGKPAAIQYGTGSIAGFFSEDSVTVGD----- 165

Qy    174 LVNIATIFESENF-----LPGI-----KWNGILGLAYATLAKPSSSLETFDLSLVTQANI 224
      :  :  :  |      ||:      |::|||  |  :  :  :
Db    166 -----LVVKDQEFIEATKEPGLTFMVAKF DGILGLGFQEISVGDA-----V 206

```

Qy	225	PNVFSMQMCG-AGLPVAGSGTN-----GGSLVLGGIEPSLYKGDIWYTPIKEEWYYQI	276
		:	
Db	207	PVWYKMVEQGLVSEPVFSFWFNRRHSDEGEGGEIVFGGMDPSHYKGNHTYVPVSQKGYWQF	266
Qy	277	EILKLEIGGQSLNLDCREYNADKAIVDSGTLLRLPQKVFDAVVEAVARASLIPE-----	331
		: : : : :	
Db	267	EMGDVLIGGKTTGF-CA--SGCSAIADSGTSLLAGPTAIITEINEKIGATGVVSQECKTV	323
Qy	332	-----FSDGF-----	336
		:	
Db	324	VSQYGQQILDLLLAETQPSKICSQVGLCTFDGKHGVSAGIKSVVDDEAGESNGLQSGPMC	383
Qy	337	-----WTGSQLACWTNSETPWSY-----FPKISIIYLRD	364
		: : :	
Db	384	NACEMAVVWMQNQLAQNKTDLILNYINQLCDKLPSPMGESSVDCGSLASMPEISFTIGA	443
Qy	365	ENSSRSFRITILPQLYIQPMMGAGLNYECY----RFGISPSTNAL-VIGATVMEGFYVIF	419
		: : : : : :	
Db	444	K-----KFALKPEEYIL-KVGEGAAQACISGFTAMDIPPRGPLWILGDVFMGAYHTVF	496
Qy	420	DRAQKRVGFAAS	431
		:	
Db	497	DYGKMRVGFAKS	508

RESULT 12

A24608

gastricsin (EC 3.4.23.3) precursor - rat

N;Alternate names: pepsinogen C

N;Contains: pepsin A (EC 3.4.23.1) precursor

C;Species: Rattus norvegicus (Norway rat)

C;Date: 30-Jun-1988 #sequence_revision 05-Aug-1994 #text_change 18-Jun-1999

C;Accession: A33510; A24608; C22434; A05145; A61298

R;Ishihara, T.; Ichihara, Y.; Hayano, T.; Katsura, I.; Sogawa, K.; Fujii-Kuriyama, Y.; Takahashi, K.

J. Biol. Chem. 264, 10193-10199, 1989

A;Title: Primary structure and transcriptional regulation of rat pepsinogen C gene.

A;Reference number: A33510; MUID:89255508; PMID:2722863

A;Accession: A33510

A;Molecule type: DNA

A;Residues: 1-392 <ISH>

A;Cross-references: GB:M25985

R;Ichihara, Y.; Sogawa, K.; Morohashi, K.; Fujii-Kuriyama, Y.; Takahashi, K. Eur. J. Biochem. 161, 7-12, 1986

A;Title: Nucleotide sequence of a nearly full-length cDNA coding for pepsinogen of rat gastric mucosa.

A;Reference number: A24608; MUID:87054020; PMID:3780741

A;Accession: A24608

A;Molecule type: mRNA

A;Residues: 1-392 <ICH>

A;Cross-references: GB:X04644; NID:g56880; PIDN:CAA28305.1; PID:g56881

R;Ichihara, Y.; Sogawa, K.; Takahashi, K.

J. Biochem. 98, 483-492, 1985

A;Title: Isolation of human, swine, and rat prepepsinogens and calf preprochymosin, and determination of the primary structures of their NH2-terminal signal sequences.

A;Reference number: A22434; MUID:86059312; PMID:2415509

A;Accession: C22434

A;Molecule type: protein

A;Residues: 1-19,'X',21-23,'X',25-29 <IC2>

R;Arai, K.M.; Muto, N.; Tani, S.; Akahane, K.

Biochim. Biophys. Acta 788, 256-261, 1984

A;Title: The N-terminal sequence of rat pepsinogen.

A;Reference number: A05145; MUID:84257697; PMID:6743670

A;Accession: A05145

A;Molecule type: protein

A;Residues: 17-30,'Q',32-102,'A',104-108,'L',110-112 <ARA>

A;Experimental source: Wistar strain

R;Ichihara, Y.; Sogawa, K.; Takahashi, K.

J. Biochem. 92, 603-606, 1982

A;Title: Rat gastric prepepsinogen: in vitro synthesis and partial amino-terminal signal sequence.

A;Reference number: A61298; MUID:83030750; PMID:6182139

A;Accession: A61298

A;Molecule type: protein

A;Residues: 1,'XX',4-6,'X',8-9,'X',11,'X',13-14,'XXX',18-19,'X',21,'X',23,'XX',26,'X' <IC3>

C;Comment: This enzyme has more restricted specificity than pepsin A. It is the major form of pepsinogen in rat gastric mucosa.

C;Genetics:

A;Introns: 20/2; 73/3; 113/1; 152/3; 219/2; 259/2; 309/3; 342/3

A;Note: there are at least two very similar genes for gastricsin in rat

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach

F;1-16/Domain: signal sequence #status experimental <SIG>

F;17-392/Product: pepsinogen #status experimental <MAT>

F;17-62/Domain: activation peptide #status experimental <ACT>

F;94,280/Active site: Asp #status predicted

F;107-112,270-275,314-347/Disulfide bonds: #status predicted

Query Match 11.6%; Score 313; DB 1; Length 392;

Best Local Similarity 29.5%; Pred. No. 1.1e-15;

Matches 105; Conservative 56; Mismatches 139; Indels 56; Gaps 16;

Qy 92 YYLEMLIGTPPQKLQILVDTGSSNEFAV-----AGTPHSYIDTYFDTERSSTYRSKGF 143
|: |: ||||| :| ||||| | | | | : |: :||| ::|

Db 76 YFGEISIGTPPQNFLVLFDTGSSNLWVSSVYCQSEACTTHA----RFNPSKSSTYYTEGQ 131

Qy 144 DVTVKYTQGSWTGFVGEDLVTPKGFNTSFLVNIATIFESENFELPG-----IKWNGILG 198
:::| || ||| | | :|: | | ||| || ::||:|

Db 132 TFSLQYGTGSLTGFFGYDTLTV-----QSIQVPNQEFGLSEN--EPGTNFVYAQFDGIMG 184

Qy 199 LAYATLAKPSSSLETFFDLSLVTQANIPNVFSMQMCGAGLPVAGS--GTNGGSLVLGGIEP 256
||| |: | | : : | : | | || |:||| :| ||::

Db 185 LAYPGLS--SGGATTALQGMLGE---GALSQPLFGVYL---GSQQGSNGGQIVFGGVDK 235

Qy 257 SLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVF 316
:|| |:| : |: :| |:|| | || |: | : |||:|:| | :| :

Db 236 NLYTGEITWVPVTQELYWQITIDDFLIGDQASGW-CSSQGC-QGIVDTGTSLLVMPAQYL 293

Qy 317 DAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRSFRITIL 376
::: : |: : | :|: : | | :| | : :

```

Db      294 SELLQTIGAQE--GEYGEYF-----VSCDSVSS-----LPTLSFVL-----NGVQFPLS 335
Qy      377 PQLY-IQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
        | | ||      : |      ::| : :| ||| :|| | |
Db      336 PSSYIIQEDNFCMVGLESLTSESGQPLWILGDVFLRSYYAIFDMGNNKVGLATS 391

```

RESULT 13

A41443

pepsin (EC 3.4.23.-) precursor, embryonic - chicken

C;Species: Gallus gallus (chicken)

C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 21-Jul-2000

C;Accession: A41443

R;Hayashi, K.; Agata, K.; Mochii, M.; Yasugi, S.; Eguchi, G.; Mizuno, T.

J. Biochem. 103, 290-296, 1988

A;Title: Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken pepsinogen: phylogenetic relationship with prochymosin.

A;Reference number: A41443; MUID:88227903; PMID:3131317

A;Accession: A41443

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-383 <HAY>

A;Cross-references: GB:D00215; NID:g2760810; PIDN:BAA00153.1; PID:g222853

C;Superfamily: pepsin

C;Keywords: aspartic proteinase; hydrolase; protein digestion

```

Query Match          11.5%; Score 310; DB 2; Length 383;
Best Local Similarity 26.8%; Pred. No. 1.8e-15;
Matches 106; Conservative 63; Mismatches 136; Indels 90; Gaps 15;

```

```

Qy      56 HA--DGLALALEPALASPAGAAFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGS 113
        || | | : || | | | | | | | | | | | : || ||
Db      55 HAFPDVLTVVTEPLL-----NTLDM-----EYYGTISIGTPPQDFTVVFDTG 97

Qy     114 SNFAVAG----TPHSYIDTYFDTERSSTYRSKGFVDTVKYTQGSWTGFVGEDLV TIPKGF 169
        || | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      98 SNLWVPSVSCTSPACQSHQMFNPSQSSTYKSTGQNLSIHYGTGDMEGTVGCDTVTVASLM 157

Qy     170 NTSFLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLVTQANI-PNVF 228
        :| : | :| | | | :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db     158 DTNQLFGLST-SEPGQFFV-YVKFDGILGLGYPSLA--ADGITPVFDNMVNESLLEQNLF 213

Qy     229 SMQMCAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSL 288
        | : : : | :| ||| : | : | | : | : : | :| : : : | :
Db     214 SVYLSREPM-----GSMVVFGGIDESYFTGSINWIPVSYQGYWQISMDSIIVNKQEI 265

Qy     289 NLDCREYNADKAIVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNS 348
        : :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :| :|
Db     266 ACS----SGCQAIIDTGTSLVAGPASDINDIQSAVG-----ANQ 300

Qy     349 ETPWSYFPKISIIYLRDENSSRSFRITILPQLYIQPMGAGLNY-----ECY 394
        | | | | | | | | | | :| : :| :| :| :| :| :| :| :| :| :| :|
Db     301 NTYGEY-----SVNCSHILAMPDVVF--VIG-GIQYPVPALAYTEQNGQGTCM 345

Qy     395 RFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFA 429
        : | : :| : :| |||| ||| |
Db     346 SSFQNSSADLWILGDVFIRVYYSIFDRANNRVGLA 380

```

RESULT 14

KHHUD

cathepsin D (EC 3.4.23.5) precursor [validated] - human

N;Alternate names: preprocathepsin D

C;Species: Homo sapiens (man)

C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 15-Sep-2000

C;Accession: A25771; S30749; PC2066; I59236; I57716

R;Faust, P.L.; Kornfeld, S.; Chirgwin, J.M.

Proc. Natl. Acad. Sci. U.S.A. 82, 4910-4914, 1985

A;Title: Cloning and sequence analysis of cDNA for human cathepsin D.

A;Reference number: A25771; MUID:85270436; PMID:3927292

A;Accession: A25771

A;Molecule type: mRNA

A;Residues: 1-412 <FAU>

A;Cross-references: EMBL:M11233; NID:g181179; PIDN:AAB59529.1; PID:g181180

R;Westley, B.R.; May, F.E.B.

Nucleic Acids Res. 15, 3773-3786, 1987

A;Title: Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive human breast cancer cells.

A;Reference number: S30749; MUID:87231068; PMID:3588310

A;Accession: S30749

A;Molecule type: mRNA

A;Residues: 1-412 <WES>

A;Cross-references: EMBL:X05344; NID:g29677; PIDN:CAA28955.1; PID:g29678

R;May, F.E.B.; Smith, D.J.; Westley, B.R.

Gene 134, 277-282, 1993

A;Title: The human cathepsin D-encoding gene is transcribed from an estrogen-regulated and a constitutive start point.

A;Reference number: PC2066; MUID:94085791; PMID:8262386

A;Accession: PC2066

A;Molecule type: DNA

A;Residues: 1-23 <MAY>

A;Cross-references: GB:L12980; NID:g291930; PIDN:AAA16314.1; PID:g455429

A;Experimental source: MCF-7 cell

R;Cavaillès, V.; Augereau, P.; Rochefort, H.

Proc. Natl. Acad. Sci. U.S.A. 90, 203-207, 1993

A;Title: Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate only TATA-dependent transcription in breast cancer cells.

A;Reference number: I59236; MUID:93126342; PMID:8419924

A;Accession: I59236

A;Status: translation not shown; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-22 <CAV1>

A;Cross-references: GB:S52557; NID:g263124; PIDN:AAD13868.1; PID:g4261568

R;Augereau, P.; Miralles, F.; Cavaillès, V.; Gaudelot, C.; Parker, M.; Rochefort, H.

Mol. Endocrinol. 8, 693-703, 1994

A;Title: Characterization of the proximal estrogen-responsive element of human cathepsin D gene.

A;Reference number: I57716; MUID:95021301; PMID:7935485

A;Accession: I57716

A;Status: translation not shown; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-22 <CAV2>

A;Cross-references: GB:S74689; NID:g786350; PIDN:AAD14156.1; PID:g4261856

R;Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A;Reference number: A51839; PDB:1LYA
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 65-161;170-241
R;Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A;Reference number: A51840; PDB:1LYB
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, with inhibitor residues 65-161;170-241
R;Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Hosur, M.V.; Sowder II, R.C.; Cachau, R.E.; Collins, J.; Silva, A.M.; Erickson, J.W.
Proc. Natl. Acad. Sci. U.S.A. 90, 6796-6800, 1993
A;Title: Crystal structures of native and inhibited forms of human cathepsin D: implications for lysosomal targeting and drug design.
A;Reference number: A48229; MUID:93342076; PMID:8393577
A;Contents: annotation; X-ray crystallography, 2.5 angstroms
C;Comment: Cathepsin D is a ubiquitous lysosomal proteinase.
C;Comment: In addition to the propeptide, residues 163-168 and 411-412 are proteolytically removed. Residues 169 and 170 are also partially removed.
C;Comment: The carbohydrate bound to 134-Asn contains a mannose-6-phosphate that is bound near 267-Lys and the phosphotransferase recognition region.
C;Genetics:
A;Gene: GDB:CTSD
A;Cross-references: GDB:120512; OMIM:116840
A;Map position: 11p15.5-11p15.5
C;Function:
A;Description: limited specificity endopeptidase
A;Pathway: intracellular protein degradation
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-64/Domain: propeptide #status predicted <PRO>
F;65-162,169-410/Product: cathepsin D #status experimental <MAT>
F;267,329-356/Region: phosphotransferase recognition
F;91-160,110-117,286-290,329-366/Disulfide bonds: #status experimental
F;97,295/Active site: Asp #status experimental
F;134,263/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 11.5%; Score 308.5; DB 1; Length 412;
Best Local Similarity 27.1%; Pred. No. 2.5e-15;
Matches 121; Conservative 75; Mismatches 180; Indels 71; Gaps 22;

Qy	9	LLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPG-----PGTPAERHADGLAL	62
		: : :: :: :	
Db	6	LLPLAL--CLLAAP--ASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPVSKYSQAVPA	61
Qy	63	ALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTP	122
		: :: : ::	
Db	62	VTEGPI--PEVLKNYM-----DAQYYGEIGIGTPPQCFTVVFDTGSSNLWVPSIH	109
Qy	123	HSYIDT-----YFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIP--KGFNTSFL	174
		: :::: : : :: : :: :	
Db	110	CKLLDIACWIHHKYNSDKSSTYVKNGTSEFDIHYGSGSLSGYLSQDTVSVPCQSASSASAL	169
Qy	175	--VNIATIFESENFFLPGI-----KWNIGILGLAYATLAKPSSSLETFFDSLVTQANI-PN	226

```

      | :      |      |||      |::|||:| | : :      ||:| : | : |
Db      170 GGVKVERQVFEATKQPGITFIAAKFDGILGMAYPRIS--VNNVLPVFDNLMQQKLVDQN 227

Qy      227 VFQMCMGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQ 286
      :||      |      || |::| | :      ||| : | : : |:: : : :|:
Db      228 IFSFY-----LSRDPDAQPGGELMLGGTDSKYYKGSLSYLVNTRKAYWQVHLDQVEV-AS 281

Qy      287 SLNLDCREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWT 346
      | | |::| : |||:|::|: |      | | |      :| : : : |
Db      282 GLTL-CKE--GCEAIVDTGTSLMVGP---VDEVRELQKAIGAVPLIQGEY----MIPC-- 329

Qy      347 NSETPWSYFPKISIIYLRDENSRSFRITILPOLYIQPMGAGLNYECYRF---GISPSTN 403
      | | |::| : :::: |: | : ||      | | | :
Db      330 ---EKVSTLPAITLKL----GGKGYKLS--PEDYTLKVSQAGKTLCLSGFMGMDIPPPSG 380

Qy      404 AL-VIGATVMEGFYVIFDRAQKRVGFA 429
      | ::| : : | :||| ||||
Db      381 PLWILGDVFIGRYYTVFDRDNNRVGFA 407

```

RESULT 15

KHMSD

cathepsin D (EC 3.4.23.5) precursor - mouse

C;Species: Mus musculus (house mouse)

C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999

C;Accession: I48278; S14704; S12587

R;Hetman, M.; Perschl, A.; Saftig, P.; Von Figura, K.; Peters, C.

DNA Cell Biol. 13, 419-427, 1994

A;Title: Mouse cathepsin D gene: molecular organization, characterization of the promoter, and chromosomal localization.

A;Reference number: I48278; MUID:94280622; PMID:8011168

A;Accession: I48278

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-410 <RES>

A;Cross-references: EMBL:X68378; NID:g50302; PIDN:CAA48453.1; PID:g817945

R;Diedrich, J.F.; Staskus, K.A.; Retzel, E.F.; Haase, A.T.

Nucleic Acids Res. 18, 7184, 1990

A;Title: Nucleotide sequence of a cDNA encoding mouse cathepsin D.

A;Reference number: S14704; MUID:91088345; PMID:2263503

A;Accession: S14704

A;Molecule type: mRNA

A;Residues: 1-410 <DIE>

A;Cross-references: EMBL:X53337; NID:g50300; PIDN:CAA37423.1; PID:g50301

R;Grusby, M.J.; Mitchell, S.C.; Glimcher, L.H.

Nucleic Acids Res. 18, 4008, 1990

A;Title: Molecular cloning of mouse cathepsin D.

A;Reference number: S12587; MUID:90326544; PMID:2374732

A;Accession: S12587

A;Molecule type: mRNA

A;Residues: 1-410 <GRU>

A;Cross-references: EMBL:X52886; NID:g50298; PIDN:CAA37067.1; PID:g50299

C;Genetics:

A;Introns: 23/2; 76/3; 118/1; 157/3; 233/2; 274/2; 322/3; 355/3

C;Function:

A;Description: limited specificity endopeptidase

A;Pathway: intracellular protein degradation

C;Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradation

```
F;1-20/Domain: signal sequence #status predicted <SIG>
```

F;21-64/Domain: propeptide #status predicted <PRO>

F;65-410/Product: cathepsin D, single-chain form #status predicted <MAT>

F;91-160,110-117,284-288,327-364/Disulfide bonds: #status predicted

F;97,293/Active site: Asp #status predicted

F;134,261/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.4%; Score 306.5; DB 1; Length 410;

Best Local Similarity 27.5%; Pred. No. 3.6e-15;

Matches 103; Conservative 64; Mismatches 123; Indels 85; Gaps 15;

Qy 92 YYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDT-----YFDTERSSTYRSKGFDV 145
|| :: ||||| :: ||||| | : ::::||| |

Db 79 YYGDIGIGITPPQCFTVVFDTGSSNLWVPSIHCKILDIACWVHHKYNSDKSSTYVKNGTSTF 138

Qy 146 TVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIAT---IFESENFFLPGI-----KWNIGIL 197
: | || :||: | :|| : | || | ||| | :|||

Db 139 DIHYGSGSLSGYLSQDTVSVPCSDQSKARGIKVEKQIF-GEATKQPGIVFVAAKFDGIL 197

Qy 198 GLAYATLAKPSSSLETFDSLVTQANI-PNVFSMQMCGAGLPVAGSGTNGGSLVLGGIEP 256
|:|::: ||:|:| : |:| | | |:| | :

Db 198 GMGYPHIS--VNNVLPVFDNLMQQKLVDKNIFSFY-----LNRDPEGQPGGELMLGGTDS 250

Qy 257 SLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVF 316
| |:: | : : |:: : : ||: | : | | : : |||::||:| | : :

Db 251 KYYHGELSYLNVTRKAYWQVHMDQLEVGNE-LTL-CK--GGCEAIVDTGTSLLVGPVEEV 306

Qy · 317 DAVVEAVARASLIPEFSDGFWTGSQLACWTNSETPWSYFPKISIIYLRDENSSRSFRITIL 376
 : :|: || : | :: |

Db 307 KELQKAIGAVPLI-----QGEYMI PCEKVSSL 333

Qy 377 PQLYIQPMMGAGLNYEC----YRFGIS-----PSTNALVIGATVMEG 414
| :|:| :| | || | :| | :| :

Db 334 PTVYLK--LG-GKNYELHPDKYILKVSQGGKTICLSGFMGMDIPPPSGPLWILGDVFIGS 390

Qy 415 FYVIFDRAQKRVGFA 429
:| :||| |||||

Db 391 YYTVFDRDNNRVGFA 405

Search completed: March 4, 2004, 15:40:59

Job time : 29.1043 secs

OM protein - protein search, using sw model

Run on: March 4, 2004, 15:39:01 ; Search time 57.8617 Seconds
(without alignments)
1890.324 Million cell updates/sec

Title: US-09-668-314C-2
Perfect score: 2687
Sequence: 1 MGALARALLLPLLAQWLLRA.....RPRDPEVVNDESSLVRHRWK 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 809742 seqs, 211153259 residues

Total number of hits satisfying chosen parameters: 809742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result		Query				Description
No.	Score	Match	Length	DB	ID	

1	2687	100.0	518	9	US-09-794-927-2	Sequence 2, Appli
2	2687	100.0	518	9	US-09-795-847-2	Sequence 2, Appli
3	2687	100.0	518	9	US-09-794-743-2	Sequence 2, Appli
4	2687	100.0	518	9	US-09-794-748-2	Sequence 2, Appli
5	2687	100.0	518	9	US-09-794-925-2	Sequence 2, Appli
6	2687	100.0	518	9	US-09-215-450-19	Sequence 19, Appl
7	2687	100.0	518	9	US-09-681-442-2	Sequence 2, Appli
8	2687	100.0	518	9	US-09-978-295A-196	Sequence 196, App
9	2687	100.0	518	9	US-09-886-143-2	Sequence 2, Appli
10	2687	100.0	518	9	US-09-978-697-196	Sequence 196, App
11	2687	100.0	518	9	US-09-978-192A-196	Sequence 196, App
12	2687	100.0	518	9	US-09-999-832A-196	Sequence 196, App
13	2687	100.0	518	10	US-09-978-189-196	Sequence 196, App
14	2687	100.0	518	10	US-09-978-608A-196	Sequence 196, App
15	2687	100.0	518	10	US-09-978-585A-196	Sequence 196, App
16	2687	100.0	518	10	US-09-978-191A-196	Sequence 196, App
17	2687	100.0	518	10	US-09-978-403A-196	Sequence 196, App
18	2687	100.0	518	10	US-09-978-564A-196	Sequence 196, App
19	2687	100.0	518	10	US-09-999-833A-196	Sequence 196, App
20	2687	100.0	518	10	US-09-981-915A-196	Sequence 196, App
21	2687	100.0	518	10	US-09-978-824-196	Sequence 196, App
22	2687	100.0	518	10	US-09-918-585A-196	Sequence 196, App
23	2687	100.0	518	10	US-09-978-423A-196	Sequence 196, App
24	2687	100.0	518	10	US-09-978-193A-196	Sequence 196, App
25	2687	100.0	518	10	US-09-869-414-2	Sequence 2, Appli
26	2687	100.0	518	10	US-09-999-830A-196	Sequence 196, App
27	2687	100.0	518	10	US-09-978-757A-196	Sequence 196, App
28	2687	100.0	518	10	US-09-978-187B-196	Sequence 196, App
29	2687	100.0	518	10	US-09-548-366-2	Sequence 2, Appli
30	2687	100.0	518	10	US-09-978-643A-196	Sequence 196, App
31	2687	100.0	518	10	US-09-978-375A-196	Sequence 196, App
32	2687	100.0	518	10	US-09-978-298A-196	Sequence 196, App
33	2687	100.0	518	10	US-09-978-188A-196	Sequence 196, App
34	2687	100.0	518	10	US-09-978-681A-196	Sequence 196, App
35	2687	100.0	518	10	US-09-978-194A-196	Sequence 196, App
36	2687	100.0	518	10	US-09-999-829A-196	Sequence 196, App
37	2687	100.0	518	10	US-09-978-299A-196	Sequence 196, App
38	2687	100.0	518	10	US-09-978-544A-196	Sequence 196, App
39	2687	100.0	518	10	US-09-978-665A-196	Sequence 196, App
40	2687	100.0	518	10	US-09-978-802A-196	Sequence 196, App
41	2687	100.0	518	13	US-10-052-586-72	Sequence 72, Appl
42	2687	100.0	518	14	US-10-174-590-72	Sequence 72, Appl
43	2687	100.0	518	14	US-10-176-758-72	Sequence 72, Appl
44	2687	100.0	518	14	US-10-175-737-72	Sequence 72, Appl
45	2687	100.0	518	14	US-10-173-706-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
 US-09-794-927-2
 ; Sequence 2, Application US/09794927
 ; Patent No. US20010016324A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Gurney, Mark E.

```

; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR,
AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/794,927
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-927-2

```

```

Query Match          100.0%; Score 2687; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-240;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy     61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGNTSFLVNIATI 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGNTSFLVNIATI 180

Qy    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA 300

Qy    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360

```

Qy 361 YLRDENSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIED 420
 |||
 Db 361 YLRDENSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIED 420

Qy 421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 |||
 Db 421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518
 |||
 Db 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518

RESULT 2

US-09-795-847-2

; Sequence 2, Application US/09795847

; Patent No. US20010018208A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.

; APPLICANT: Parodi, Luis A.

; APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR,
 AND

; TITLE OF INVENTION: USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 28341/6280DE

; CURRENT APPLICATION NUMBER: US/09/795,847

; CURRENT FILING DATE: 2001-02-28

; PRIOR APPLICATION NUMBER: 09/416,901

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 2

; LENGTH: 518

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-795-847-2

Query Match 100.0%; Score 2687; DB 9; Length 518;

Best Local Similarity 100.0%; Pred. No. 2.8e-240;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
 |||
 Db 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy 61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Db	61	ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Qy	361	YLRDENS SRSFRTILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENS SRSFRTILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLV LIVLLLLPFCRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLV LIVLLLLPFCRRPRDPEVVNDESSLVRHRWK	518

RESULT 3

US-09-794-743-2

; Sequence 2, Application US/09794743

; Patent No. US20010021391A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.

; APPLICANT: Parodi, Luis A.

; APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR,
AND

; TITLE OF INVENTION: USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 28341/6280BC

; CURRENT APPLICATION NUMBER: US/09/794,743

; CURRENT FILING DATE: 2001-02-27

; PRIOR APPLICATION NUMBER: 09/416,901

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-743-2

Query Match 100.0%; Score 2687; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-240;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGP	60
Qy	61	ALALEPALASPAGAAANFLAMVDNLQD	120
Db	61	ALALEPALASPAGAAANFLAMVDNLQD	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGF	180
Db	121	TPHSYIDTYFDTERSSTYRSKGF	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGD	300
Db	241	GSGTNGGSLVLGGIEPSLYKGD	300
Qy	301	IVDSGTLLRLPQKVFD	360
Db	301	IVDSGTLLRLPQKVFD	360
Qy	361	YLRDENS	420
Db	361	YLRDENS	420
Qy	421	RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLV	518
Db	481	AILLV	518

RESULT 4

US-09-794-748-2

; Sequence 2, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.

```

; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR,
AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280JL
; CURRENT APPLICATION NUMBER: US/09/794,748
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-748-2

```

```

Query Match          100.0%;  Score 2687;  DB 9;  Length 518;
Best Local Similarity 100.0%;  Pred. No. 2.8e-240;
Matches 518;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      1 MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGENTSFLVNIATI 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGENTSFLVNIATI 180

Qy    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300

Qy    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFS DGFWTGSQ LACWTNSETPWSYFPKISI 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFS DGFWTGSQ LACWTNSETPWSYFPKISI 360

```

Qy	361	YLRDENSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK	518

Query Match 100.0%; Score 2687; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-240;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Db          61 ALALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
Qy          121 TPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
            |||
Db          121 TPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
Qy          181 FESEFFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
            |||
Db          181 FESEFFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
Qy          241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
            |||
Db          241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
Qy          301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
            |||
Db          301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
Qy          361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
            |||
Db          361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
Qy          421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
            |||
Db          421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
Qy          481 AILLVLIVLLLLPFRQCRRPRDPEVVNDESSLVRHRWK 518
            |||
Db          481 AILLVLIVLLLLPFRQCRRPRDPEVVNDESSLVRHRWK 518

```

RESULT 6

US-09-215-450-19

```

; Sequence 19, Application US/09215450
; Patent No. US20020068278A1
; GENERAL INFORMATION:
; APPLICANT: Giese, Klaus
; APPLICANT: Xin, Hong
; TITLE OF INVENTION: METASTATIC BREAST AND COLON CANCER REGULATED GENES
; FILE REFERENCE: 1451.100 / 210030.447
; CURRENT APPLICATION NUMBER: US/09/215,450
; CURRENT FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
; LENGTH: 518
; TYPE: PRT
; ORGANISM: human

```

US-09-215-450-19

```

Query Match          100.0%;  Score 2687;  DB 9;  Length 518;
Best Local Similarity 100.0%;  Pred. No. 2.8e-240;
Matches 518;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy          1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
            |||
Db          1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

```

Qy 61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
 |||||
 Db 61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy 121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGNTSFLVNIATI 180
 |||||
 Db 121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGNTSFLVNIATI 180

Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 |||||
 Db 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 |||||
 Db 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300

Qy 301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 |||||
 Db 301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360

Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 |||||
 Db 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

Qy 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 |||||
 Db 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy 481 AILLVLIVLLLLPFCRQRPRDPEVVNDESSLVRHRWK 518
 |||||
 Db 481 AILLVLIVLLLLPFCRQRPRDPEVVNDESSLVRHRWK 518

RESULT 7

US-09-681-442-2

; Sequence 2, Application US/09681442

; Patent No. US20020081634A1

; GENERAL INFORMATION:

; APPLICANT: Gurney, Mark E.

; APPLICANT: Bienkowski, Michael J.

; APPLICANT: Heinrikson, Robert L.

; APPLICANT: Parodi, Luis A.

; APPLICANT: Yan, Riqiang

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND USES

; TITLE OF INVENTION: THEREFOR

; FILE REFERENCE: 28341/6280FG

; CURRENT APPLICATION NUMBER: US/09/681,442

; CURRENT FILING DATE: 2001-04-05

; PRIOR APPLICATION NUMBER: 09/416,901

; PRIOR FILING DATE: 1999-10-13

; PRIOR APPLICATION NUMBER: 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-681-442-2

Query Match 100.0%; Score 2687; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-240;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPT	PGPGTPAERHADGL	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPT	PGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAAANFLAMVDNLQGD	SGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAAANFLAMVDNLQGD	SGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDVT	VKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDVT	VKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSS	LETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSS	LETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGD	IWYTPIKEEWYYQIEILKLEIGGQSLNLD	CREYNADKA 300
Db	241	GSGTNGGSLVLGGIEPSLYKGD	IWYTPIKEEWYYQIEILKLEIGGQSLNLD	CREYNADKA 300
Qy	301	IVDSGTTLLRLPQKVFD	AVVEAVARASLIPEFSDGFWTGSQ	LACWTNSETPWSYFPKISI 360
Db	301	IVDSGTTLLRLPQKVFD	AVVEAVARASLIPEFSDGFWTGSQ	LACWTNSETPWSYFPKISI 360
Qy	361	YLRDENSSRSFRITILPQ	LYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSSRSFRITILPQ	LYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAEIAGA	AAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAEIAGA	AAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLVLIIVLLLLP	FRCQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLVLIIVLLLLP	FRCQRRPRDPEVVNDESSLVRHRWK	518

RESULT 8
US-09-978-295A-196
; Sequence 196, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C11
; CURRENT APPLICATION NUMBER: US/09/978,295A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936

; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15

; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600

; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 100.0%; Score 2687; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-240;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGPGPCTPAERHADGL 60
|
Db 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGPGPCTPAERHADGL 60

Qy 61 ALALEPALASPAGAANFLAMVDNLQDGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
|
Db 61 ALALEPALASPAGAANFLAMVDNLQDGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy 121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
|
Db 121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180

Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
|
Db 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA 300
|
Db 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA 300

Qy 301 IVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360
|
Db 301 IVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360

Qy 361 YLRDENSSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
|

```

Db          361 YLRDENSRSRFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
Qy          421 RAQKRVGFAASPCAETIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
          |||
Db          421 RAQKRVGFAASPCAETIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
Qy          481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518
          |||
Db          481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518

```

Query Match 100.0%; Score 2687; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-240;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 Qy 301 IVDSGTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 Db 301 IVDSGTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 Db 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 Qy 421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 Db 421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 Qy 481 AILLVLIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK 518
 Db 481 AILLVLIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK 518

RESULT 10

US-09-978-697-196

; Sequence 196, Application US/09978697

; Patent No. US20020169284A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan

; APPLICANT: Ferrara, Napoleon

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Kuo, Sophia S.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James;

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Shelton, David L.

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2630P1C27

; CURRENT APPLICATION NUMBER: US/09/978,697

; CURRENT FILING DATE: 2001-10-16

; PRIOR APPLICATION NUMBER: 09/918585

; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01

; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554

```
Query Match      100.0%;   Score 2687;   DB 9;   Length 518;
Best Local Similarity 100.0%;   Pred. No. 2.8e-240;
Matches 518;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;
```

Qy 1 MGALARALLLPLLAQWLLRAAPELAPAFPTLPLRVAAATNRVVAPTGPGTPAERHADGL 60
 |||||

Db 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGPGPCTPAERHADGL 60

Qy 61 ALALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
 |||

Db 61 ALALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy 121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
 |||

Db 121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180

Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 |||

Db 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240

Qy 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 |||

Db 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300

Qy 301 IVDSGTTLLRLPQKVFD AVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 |||

Db 301 IVDSGTTLLRLPQKVFD AVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360

Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 |||

Db 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

Qy 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 |||

Db 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy 481 AILLVLIVL LLLPFR CQRRPRDPEVVNDESSLVRHRWK 518
 |||

Db 481 AILLVLIVL LLLPFR CQRRPRDPEVVNDESSLVRHRWK 518

RESULT 11

US-09-978-192A-196

; Sequence 196, Application US/09978192A

; Patent No. US20020177553A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J
 ; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C9
; CURRENT APPLICATION NUMBER: US/09/978,192A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27

; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796

; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15

; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 100.0%; Score 2687; DB 9; Length 518;
Best Local Similarity 100.0%; Pred. No. 2.8e-240;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
|
Db 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy 61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
|
Db 61 ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy 121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTPKGFNTSFLVNIATI 180
|
Db 121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTPKGFNTSFLVNIATI 180

Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLVTQANIPNVFSMQMCGAGLPVA 240
|
Db 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLVTQANIPNVFSMQMCGAGLPVA 240

Qy 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA 300
|
Db 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA 300

Qy 301 IVDSGTTLLRLPQKVFDAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360
|
Db 301 IVDSGTTLLRLPQKVFDAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360

Qy 361 YLRDENSSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
|
Db 361 YLRDENSSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

Qy 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
|
Db 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518
|
Db 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518

RESULT 12

US-09-999-832A-196

; Sequence 196, Application US/09999832A

; Publication No. US20020192706A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C63
; CURRENT APPLICATION NUMBER: US/09/999,832A
; CURRENT FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20

; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952

; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07

```

; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

```

Query Match          100.0%;  Score 2687;  DB 9;  Length 518;
Best Local Similarity 100.0%;  Pred. No. 2.8e-240;
Matches 518;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFGVGEDLVTIPKGFNTSFLVNIATI 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFGVGEDLVTIPKGFNTSFLVNIATI 180

Qy    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLVTQANIPNVFSMQMCGAGLPVA 240
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFDLSLVTQANIPNVFSMQMCGAGLPVA 240

Qy    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA 300
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA 300

Qy    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360

Qy    361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420

```

Qy 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 |||
 Db 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480

Qy 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518
 |||
 Db 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518

RESULT 13

US-09-978-189-196

; Sequence 196, Application US/09978189

; Publication No. US20030004102A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James;
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; TITLE OF INVENTION: Acids Encoding the Same

; FILE REFERENCE: P2630P1C7

; CURRENT APPLICATION NUMBER: US/09/978,189

; CURRENT FILING DATE: 2001-10-15

; PRIOR APPLICATION NUMBER: 09/918585

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/064249

; PRIOR FILING DATE: 1997-11-03

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066364

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049

; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30

```

; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

```

```

Query Match          100.0%;  Score 2687;  DB 10;  Length 518;
Best Local Similarity 100.0%;  Pred. No. 2.8e-240;
Matches 518;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

```

```

Qy      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60

Qy     61 ALALEPALASPAGAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ALALEPALASPAGAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120

Qy    121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
        ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180

```

Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFD	SLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFD	SLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEE	WYQIEILKLEIGGQSLNLDCREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEE	WYQIEILKLEIGGQSLNLDCREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFD	AVVEAVARASLIPEFSDGFWTGSQ	LACWTNSETPWSYFPKISI 360
Db	301	IVDSGTTLLRLPQKVFD	AVVEAVARASLIPEFSDGFWTGSQ	LACWTNSETPWSYFPKISI 360
Qy	361	YLRDENSSRSFRITILPQLYIQPMMGAGL	NYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSSRSFRITILPQLYIQPMMGAGL	NYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAEIAGAAVSEISGPF	STEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAEIAGAAVSEISGPF	STEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLVLIVLLLLPFR	CQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLVLIVLLLLPFR	CQRRPRDPEVVNDESSLVRHRWK	518

RESULT 14

US-09-978-608A-196

; Sequence 196, Application US/09978608A

; Publication No. US20030045462A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J
 ; APPLICANT: Kljavin, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James;
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2630P1C22
 ; CURRENT APPLICATION NUMBER: US/09/978,608A
 ; CURRENT FILING DATE: 2001-10-16
 ; NUMBER OF SEQ ID NOS: 624
 ; Prior Application removed - See File Wrapper or Palm
 ; SEQ ID NO 196
 ; LENGTH: 518
 ; TYPE: PRT
 ; ORGANISM: Homo sapien
 US-09-978-608A-196

Query Match 100.0%; Score 2687; DB 10; Length 518;
 Best Local Similarity 100.0%; Pred. No. 2.8e-240;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGGPGTPAERHADGL	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GS GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Db	241	GS GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSRSFRITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Qy	481	AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK	518
Db	481	AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK	518

RESULT 15
 US-09-978-585A-196

Db 121 TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGENTSFLVNIATI 180
 Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 Qy 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 241 GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 Qy 301 IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 301 IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 Qy 421 RAQKRVGFAASPCA EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 421 RAQKRVGFAASPCA EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 Qy 481 AILLVLIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK 518
 ||||||||||||||||||||||||||||||||||||||||||||
 Db 481 AILLVLIVLLLLPFR CQRRPRDPEVVNDESSLVRHRWK 518

Search completed: March 4, 2004, 15:57:36
 Job time : 59.8617 secs

OM protein - protein search, using sw model

Run on: March 4, 2004, 15:28:35 ; Search time 74.3936 Seconds
(without alignments)
2196.942 Million cell updates/sec

Title: US-09-668-314C-2
Perfect score: 2687
Sequence: 1 MGALARALLLPLLAQWLLRA.....RPRDPEVVNDESSLVRHRWK 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_25:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

%
Result Query
No. Score Match Length DB ID Description

1	2405	89.5	514	11	Q8C5E9	Q8c5e9 mus musculu
2	2399	89.3	514	11	Q8C793	Q8c793 mus musculu
3	2395	89.1	514	11	Q9JL18	Q9jl18 mus musculu
4	2375	88.4	468	4	Q9NZL2	Q9nzl2 homo sapien
5	2293	85.3	439	4	Q9H2V8	Q9h2v8 homo sapien
6	2184	81.3	423	4	Q8N2D4	Q8n2d4 homo sapien
7	1966	73.2	396	4	Q9NZL1	Q9nzl1 homo sapien
8	1693	63.0	500	13	Q7T0Y2	Q7t0y2 xenopus lae
9	1246	46.4	255	11	Q9R1P7	Q9rlp7 mus musculu
10	1186	44.1	501	11	Q8C7R1	Q8c7r1 mus musculu
11	1183.5	44.0	532	4	Q9ULS1	Q9uls1 homo sapien
12	1183	44.0	501	11	Q8BQY4	Q8bqy4 mus musculu
13	1172.5	43.6	501	4	Q8IYC8	Q8iyc8 homo sapien
14	1049	39.0	467	11	Q8C4F4	Q8c4f4 mus musculu
15	653	24.3	267	11	Q9CUU5	Q9cuu5 mus musculu
16	461	17.2	213	4	Q9P0D2	Q9p0d2 homo sapien
17	386	14.4	244	5	Q8WQY9	Q8wqy9 aphrocallis
18	367.5	13.7	383	13	Q9DEC3	Q9dec3 xenopus lae
19	361.5	13.5	389	6	Q9GMY4	Q9gmy4 sorex ungui
20	355.5	13.2	384	13	Q91322	Q91322 rana catesb
21	355	13.2	389	13	Q9W643	Q9w643 gallus gall
22	355	13.2	389	13	Q9PWK1	Q9pwk1 gallus gall
23	351.5	13.1	389	6	Q9GMY3	Q9gmy3 rhinolophus
24	351.5	13.1	389	6	Q9GMY5	Q9gmy5 suncus muri
25	335.5	12.5	388	6	Q9GMY2	Q9gmy2 oryctolagus
26	334.5	12.4	391	5	Q9VKP6	Q9vvp6 drosophila
27	326	12.1	399	13	O93458	O93458 podarcis si
28	324.5	12.1	372	5	Q9VLK3	Q9vlk3 drosophila
29	324.5	12.1	383	13	Q9DE45	Q9de45 salvelinus
30	322.5	12.0	390	6	Q8SQ41	Q8sq41 canis famil
31	322.5	12.0	397	13	Q9W6D4	Q9w6d4 hynobius le
32	320	11.9	385	13	Q9DEC4	Q9dec4 rana catesb
33	319.5	11.9	387	13	Q9DDV5	Q9ddv5 salvelinus
34	319	11.9	419	5	Q95VA2	Q95va2 clonorchis
35	318.5	11.9	396	13	O93428	O93428 chionodraco
36	318	11.8	397	13	Q800A0	Q800a0 rana catesb
37	316	11.8	378	13	Q9PUR9	Q9pur9 pseudopleur
38	316	11.8	392	11	Q9D7R7	Q9d7r7 mus musculu
39	313	11.6	383	5	O76856	O76856 dictyosteli
40	312.5	11.6	354	5	Q9GYX7	Q9gyx7 boophilus m
41	305	11.4	384	13	Q9DEC2	Q9dec2 xenopus lae
42	302	11.2	398	13	P87370	P87370 oncorhynch
43	301.5	11.2	401	11	Q91X66	Q91x66 mus musculu
44	300.5	11.2	386	6	Q9BGU5	Q9bgu5 bos taurus
45	295.5	11.0	390	6	Q9GK10	Q9gk10 camelus dro

ALIGNMENTS

RESULT 1

Q8C5E9

ID Q8C5E9 PRELIMINARY; PRT; 514 AA.
AC Q8C5E9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

Db 357 YLRDENASRSFRITILPQLYIQPMMGAGFNYECYRFGISSSTNALVIGATVMEGFYVVF 416

Qy 421 RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 |||:||||| ||||| | |||||:|||||:|:|||||

Db 417 RAQRRVGFAVSPCAIEIGTTVSEISGPFSTEDIASNCVPAQALNEPILWIVSYALMSVCG 476

Qy 481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518
 |||||:|||| |: |||||

Db 477 AILLVLILLLLLPLHCRHAPRDPEVVNDESSLVRHRWK 514

RESULT 2

Q8C793

ID Q8C793 PRELIMINARY; PRT; 514 AA.

AC Q8C793;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Beta-site APP-cleaving enzyme 2.

GN BACE2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Heart;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR EMBL; AK052309; BAC34931.1; -.

DR MGD; MGI:1860440; Bace2.

DR GO; GO:0004194; F:pepsin A activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001969; Aspprotease_AS.

DR InterPro; IPR001461; Peptidase_A1.

DR InterPro; IPR009007; Pept_A_acid.

DR Pfam; PF00026; asp; 1.

DR PRINTS; PR00792; PEPSIN.

DR PROSITE; PS00141; ASP_PROTEASE; 2.

SQ SEQUENCE 514 AA; 55871 MW; 8BF45E07B0990225 CRC64;

Query Match 89.3%; Score 2399; DB 11; Length 514;

Best Local Similarity 88.6%; Pred. No. 1e-178;

Matches 459; Conservative : 20; Mismatches 35; Indels 4; Gaps 1;

Qy 1 MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL 60
 |||| |||||:||||| | | |||||:|| ||| : || ||| ||||

Db 1 MGALLRALLLPVLAQWLLSAVPALAPAPFTLPLQVARATNHRASAVPGLGTPELPRADGL 60

Qy 61 ALALEPALASPAGAANFLAMVDNLQGDSSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG 120
 ||||| |: |||||:|||||

Db 61 ALALEPVRAT----ANFLAMVDNLQGDSSGRGYYLEMLIGTPPQKVQILVDTGSSNFAVAG 116

Qy 121 TPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI 180
 |||||:| ||| |||||:|||||
 Db 117 APHSYIDTYFDSESSSTYHSKGFDTVVKYTQGSWTGFVGEDLVTIPKGFNSSFLVNIATI 176

Qy 181 FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA 240
 |||||:| ||| |||||:| |||:| |||||
 Db 177 FESENFFLPGIKWNGILGLAYAALAKPSSSLETFFDSLVAQAKIPDIFSMQMCGAGLPVA 236

Qy 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 |||||:| ||| |||||:| |||:| |||||
 Db 237 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 296

Qy 301 IVDSGTTLLRLPQKVFDAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
 |||||:| ||| |||||:| |||:| |||||
 Db 297 IVDSGTTLLRLPQKVFDAVVEAVARTSLIPEFSDGFWTGAQLACWTNSETPWAYFPKISI 356

Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 420
 |||||:| ||| |||||:| |||:| |||||:| |||
 Db 357 YLRDENASRSFRTTILPQLYIQPMMGAGFN YECYRFGISSSTNALVIGATVMEGFYVVFD 416

Qy 421 RAQKRVGFAASPCAEIAGA AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 |||:| ||| ||||| | |||||:| |||||:| |||||:| |||||
 Db 417 RAQRRVGFAVSPCAEIEGTTVSEISGPFSTEDIASNCVPAQALNEPILWIVSYALMSVCG 476

Qy 481 AILLVLIVLLLLL PFCQRRPRDPEVVNDESSLVRHRWK 518
 |||||:| ||| | : |||||
 Db 477 AILLVLILLLLLL PLHCRHAPRDPEVVNDESSLVRHRWK 514

RESULT 3

Q9JL18

ID Q9JL18 PRELIMINARY; PRT; 514 AA.
 AC Q9JL18;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Aspartyl protease 1.
 GN BACE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Choi D.K., Sugano S., Sakaki Y.;
 RT "Molecular characterization of the mouse Asp1 gene, a homolog of the
 RT human ASP1 (Down Syndrome Region aspartyl protease).";
 RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
 DR EMBL; AF216310; AAF36599.1; -.
 DR HSSP; P00797; 2REN.
 DR MEROPS; A01.041; -.
 DR MGD; MGI:1860440; Bace2.
 DR GO; GO:0004194; F:pepsin A activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001969; Aspprotease_AS.

DR InterPro; IPR001461; Peptidase_A1.
 DR InterPro; IPR009007; Pept_A_acid.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Aspartyl protease; Hydrolase; Protease.
 SQ SEQUENCE 514 AA; 55799 MW; A70725F2C1DF5B47 CRC64;

Query Match 89.1%; Score 2395; DB 11; Length 514;
 Best Local Similarity 88.6%; Pred. No. 2.1e-178;
 Matches 459; Conservative 20; Mismatches 35; Indels 4; Gaps 1;

Qy	1	MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
		: : :	
Db	1	MGALLRALLLLVLAQWLLSAVPALAPAPFTLPLQVAGATNHRASAVPGLGTPELPRADGL	60
Qy	61	ALALEPALASPAGAAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
		: : : :	
Db	61	ALALEPVRAT----ANFLAMVDNLQGDSGRGYYLEMLIGTPPQKVQILVDTGSSNFAVAG	116
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATI	180
		: : :	
Db	117	APHSYIDTYFDSESSSTYHSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNSFLVNIATI	176
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFDLSVTQANIPNVFSMQMCGAGLPVA	240
		: : : : :	
Db	177	FESENFFLPGIKWNGILGLAYAALAKPSSSLETFDLSVAQAKIPDIFSMQMCGAGLPVA	236
Qy	241	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA	300
		: : : :	
Db	237	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQNLNLDREYNADKA	296
Qy	301	IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
		: : : :	
Db	297	IVDSGTTLLRLPQKVFDVAVVEAVARTSLIPEFSDGFWTGAQACWTNSETPWAYFPKISI	356
Qy	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNYECYREGISPSTNALVIGATVMEGFYVIFD	420
		: : : :	
Db	357	YLRDENASRSFRITILPQLYIQPMMGAGFNYECYREGISSSTNALVIGATVMEGFYVIFD	416
Qy	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
		: : : :	
Db	417	RAQRRVGFAVSPCAIEIEGTTVSEISGPFSTEDIASNCVPAQALNEPILWIVSYALMSVCG	476
Qy	481	AILLVLIIVLILLPLFRQRRPRDPEVVNDESSLVRHRWK	518
		: :	
Db	477	AILLVLIIVLILLPLHCRHAPRDPEVVNDESSLVRHRWK	514

RESULT 4

Q9NZL2

ID Q9NZL2 PRELIMINARY; PRT; 468 AA.
 AC Q9NZL2;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Aspartyl protease.

GN BACE2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20422477; PubMed=10965118;
 RA Solans A., Estivill X., de La Luna S.;
 RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
 RT Alzheimer's amyloid precursor protein beta-secretase.";
 RL Cytogenet. Cell Genet. 89:177-184(2000).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
 DR EMBL; AF188276; AAF35835.1; -.
 DR HSSP; P00797; 2REN.
 DR GO; GO:0004194; F:pepsin A activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR001461; Peptidase_A1.
 DR InterPro; IPR009007; Pept_A_acid.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Aspartyl protease; Hydrolase; Protease.
 SQ SEQUENCE 468 AA; 50324 MW; 717E0920126A0142 CRC64;

Query Match 88.4%; Score 2375; DB 4; Length 468;
 Best Local Similarity 90.3%; Pred. No. 6.6e-177;
 Matches 468; Conservative 0; Mismatches 0; Indels 50; Gaps 1;

Qy	1	MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Db	1	MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDSGTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Db	301	IVDSGTLLRLPQKVFDVAVVEAVARASL-----	328
Qy	361	YLRDENSRSRFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420

Db 329 -----LYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD 370

Qy 421 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 480
 |||

Db 371 RAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG 430

Qy 481 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 518
 |||

Db 431 AILLVLIVLLLLPFRQRRPRDPEVVNDESSLVRHRWK 468

RESULT 5

Q9H2V8

ID Q9H2V8 PRELIMINARY; PRT; 439 AA.

AC Q9H2V8;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE CDA13.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Pheochromocytoma;

RA Li Y., Huang Q., Peng, y, Song H., Yu Y., Xu S., Ren S., Chen Z.,

RA Han Z.;

RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.

DR EMBL; AF212252; AAG41783.1; -.

DR HSSP; P00797; 2REN.

DR GO; GO:0004194; F:pepsin A activity; IEA.

DR GO; GO:0008233; F:peptidase activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001969; Aspprotease_AS.

DR InterPro; IPR001461; Peptidase_A1.

DR InterPro; IPR009007; Pept_A_acid.

DR Pfam; PF00026; asp; 1.

DR PRINTS; PR00792; PEPSIN.

DR PROSITE; PS00141; ASP_PROTEASE; 2.

KW Aspartyl protease; Hydrolase; Protease.

SQ SEQUENCE 439 AA; 48275 MW; 02EC0E0E50F11602 CRC64;

Query Match 85.3%; Score 2293; DB 4; Length 439;

Best Local Similarity 100.0%; Pred. No. 1.5e-170;

Matches 439; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 MVDNLQGDSSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDTYFDTERSSTYR 139
 |||

Db 1 MVDNLQGDSSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDTYFDTERSSTYR 60

Qy 140 SKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENFFLPGIKWNGILGL 199
 |||

Db 61 SKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENFFLPGIKWNGILGL 120

Qy 200 AYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLY 259

```

      |||
Db      121 AYATLAKPSSSLETFFDLSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLY 180

Qy      260 KGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVFDAY 319
      |||
Db      181 KGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVFDAY 240

Qy      320 VEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISIIYLRDENSSRSFRITILPQL 379
      |||
Db      241 VEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISIIYLRDENSSRSFRITILPQL 300

Qy      380 YIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCAIEIAGA 439
      |||
Db      301 YIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCAIEIAGA 360

Qy      440 AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLLPFCRR 499
      |||
Db      361 AVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLLPFCRR 420

Qy      500 PRDPEVVNDESSLVRHRWK 518
      |||
Db      421 PRDPEVVNDESSLVRHRWK 439

```

RESULT 6

Q8N2D4

```

ID   Q8N2D4          PRELIMINARY;      PRT;   423 AA.
AC   Q8N2D4;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein OVARC1000363.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Ovarian carcinoma;
RA   Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
RA   Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,
RA   Nagahari K., Sugano S., Isogai T.;
RT   "HRI human cDNA sequencing project.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK075539; BAC11682.1; -.
DR   GO; GO:0004194; F:pepsin A activity; IEA.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR   InterPro; IPR001969; Aspprotease_AS.
DR   InterPro; IPR001461; Peptidase_A1.
DR   InterPro; IPR009007; Pept_A_acid.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hypothetical protein.
SQ   SEQUENCE 423 AA; 46457 MW; 4D4839F2ED9C2CE1 CRC64;

```

Query Match 81.3%; Score 2184; DB 4; Length 423;

Best Local Similarity 99.3%; Pred. No. 4.7e-162;
Matches 420; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy      96 MLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWT 155
          |||
Db      1 MLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWT 60

Qy     156 GFVGEDLVTIPKGFNTSFLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFF 215
          |||
Db      61 GFVGEDLVTIPKGFNTSFLVNIATIFESGNFFLPGIQWNGILGLAYATLAKPSSSLETFF 120

Qy     216 DSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQ 275
          |||
Db     121 DSLVTQANIPNVFSMQMRGAGLPVAGSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQ 180

Qy     276 IEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVFDAVVEAVARASLIPEFSDG 335
          |||
Db     181 IEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVFDAVVEAVARASLIPEFSDG 240

Qy     336 FWTGSQLACWTNSETPWSYFPKISIIYLRDENSRRSFRITILPQLYIQPMMGAGLNYECYR 395
          |||
Db     241 FWTGSQLACWTNSETPWSYFPKISIIYLRDENSRRSFRITILPQLYIQPMMGAGLNYECYR 300

Qy     396 FGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVAS 455
          |||
Db     301 FGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCAEIAGAAVSEISGPFSTEDVAS 360

Qy     456 NCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLLPFRRCRRPRDPEVVNDESSLVRH 515
          |||
Db     361 NCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLLLPFRRCRRPRDPEVVNDESSLVRH 420

Qy     516 RWK 518
          |||
Db     421 RWK 423
```

RESULT 7

Q9NZL1

ID Q9NZL1 PRELIMINARY; PRT; 396 AA.
AC Q9NZL1;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Aspartyl protease.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase."
RL Cytogenet. Cell Genet. 89:177-184(2000).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.

DR EMBL; AF188277; AAF35836.1; -.
 DR HSSP; P00797; 2REN.
 DR GO; GO:0004194; F:pepsin A activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR001461; Peptidase_A1.
 DR InterPro; IPR009007; Pept_A_acid.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Aspartyl protease; Hydrolase; Protease.
 SQ SEQUENCE 396 AA; 43013 MW; 5023A7AF391CEAC9 CRC64;

Query Match 73.2%; Score 1966; DB 4; Length 396;
 Best Local Similarity 100.0%; Pred. No. 4.5e-145;
 Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Db	1	MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTGQSWTGFGEDLVTIPKGFNTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTGQSWTGFGEDLVTIPKGFNTSFLVNIATI	180
Qy	181	FESNFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESNFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSNTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA	300
Db	241	GSNTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKA	300
Qy	301	IVDSGTTLLRLPQKVFDAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLLRLPQKVFDAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSSRSFRITILPQ	378
Db	361	YLRDENSSRSFRITILPQ	378

RESULT 8

Q7T0Y2

ID Q7T0Y2 PRELIMINARY; PRT; 500 AA.
 AC Q7T0Y2;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Qy 133 ERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENFFLPGIK 192
 : |::|:| :|||:||||||| :|:|:|:| | :||:|:|:|:|:|:| |
 Db 117 KLSTSYQSLNTEVTVRYTQGSWTGLLGKDVVSIPKGVNGTFLINIASIFQSESFFLPNIN 176

Qy 193 WNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGSLVLG 252
 | |||||:|||||||:| ||||| | |||:||||||| | :| | |||||
 Db 177 WQGILGLAYSTLAKPSSSVEPFFDSLQVENIPDVFSMQMCGAGQSSPGNGINAGSLVLG 236

Qy 253 GIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLRLP 312
 |:|||||:||||| |||||:|:| |:| | |||| | |:||||| |||||
 Db 237 GVEPSLYKGNWYTPITEEWYYQVEVLKFEVGGQRLNLDCTVYNSDKAIVDSGTTLRLP 296

Qy 313 QKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSRSRFR 372
 |||:|:|:|: : ||| |: || | |||| :| ||:| | ||||| |:| ||||
 Db 297 DKVFNAMVDAIVQTSLIQNFNAEFWAGLQACWDKTQQPWNYFPDISIIYLRDTNTRSFR 356

Qy 373 ITILPQLYIQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASP 432
 :|: ||||| :| : |:|||| | : ||||| ||||| |||||:||||| |
 Db 357 LTLKPQLYIQSVLTFQESLNCFRFGISQSASTLVIGATVMEGFYVIFDRAEKRVGFAVSS 416

Qy 433 CAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLLVLLL 492
 |||::| |||:| | ||:|:| : | |||:|:|:|:|:| | ||||:| ||||
 Db 417 CAEVSGITVSEIAGPFGTSDVSSNCIARNPLREPIMWIIISYSLMSLCGMILLVLVLLL 476

Qy 493 PFCRCRRPRDPEVVNDESSLVRHRWK 518
 | :| | | :|||||:||||
 Db 477 SNR--QRHDDMETINDESSLVQHRWK 500

RESULT 9

Q9R1P7

ID Q9R1P7 PRELIMINARY; PRT; 255 AA.
 AC Q9R1P7;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Aspartyl protease (Fragment).
 GN BACE2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Accarino M., Fumagalli P., Taramelli R., Ottolenghi S.;
 RT "Cloning of a gene from chromosome 21 Down Region encoding a potential
 RT transmembrane protease."
 RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; AF051150; AAD45964.1; -.
 DR MEROPS; A01.041; -.
 DR MGD; MGI:1860440; Bace2.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; IEA.
 DR GO; GO:0008233; F:peptidase activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.

KW Protease.

FT NON_TER 1 1

SQ SEQUENCE 255 AA; 28685 MW; 53DE317815996D63 CRC64;

Query Match 46.4%; Score 1246; DB 11; Length 255;

Best Local Similarity 91.0%; Pred. No. 4e-89;

Matches 232; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

Qy 264 WYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKAI VDS GTTLLRLPQKVFD A VVEAV 323

||||| : |||||

Db 1 WYTPIKEEWYYQIEILKLEIGGQNLNLD CREYNADKAI VDS GTTLLRLPQKVFD A VVEAV 60

Qy 324 ARASLIPEFS DGF WTGS QLACWTNSET PWSYFPKIS IYLRDENS SRSF RITILPQLYIQP 383

|| ||||| : ||||| : ||||| : |||||

Db 61 ARTSLIPEFS DGF WTGAQLACWTNSET PWAYFPKIS IYLRDENAS RSF RITILPQLYIQP 120

Qy 384 MMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGF AASPCA EIAGAAVSE 443

||||| ||||| ||||| : ||||| : ||||| ||||| | |||

Db 121 MMGAGFN YECYRFGISSSTNALVIGATVMEGFYVVFDR AQRVGF A VSPCAEIEGTTVSE 180

Qy 444 ISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLV LVLV LLLL PFCRRPRDP 503

||||| : ||||| : ||||| : ||||| : ||| : |||

Db 181 ISGPFSTEDIASNCVPAQALNEPILWIVSYALMSVCGAILLV L L L L L VPLHCRHAPRDP 240

Qy 504 EVVNDESSLVRHRWK 518

|||||

Db 241 EVVNDESSLVRHRWK 255

RESULT 10

Q8C7R1

ID Q8C7R1 PRELIMINARY; PRT; 501 AA.

AC Q8C7R1;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Beta-site APP cleaving enzyme.

GN BACE.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Spinal cord;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs.";

RL Nature 420:563-573(2002).

DR EMBL; AK049626; BAC33844.1; -.

DR MGD; MGI:1346542; Bace.

DR GO; GO:0004194; F:pepsin A activity; IEA.

DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.

DR InterPro; IPR001969; Aspprotease_AS.

DR InterPro; IPR001461; Peptidase_A1.

DR InterPro; IPR009007; Pept_A_acid.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
SQ SEQUENCE 501 AA; 55761 MW; B410DA8B64647663 CRC64;

Query Match 44.1%; Score 1186; DB 11; Length 501;
Best Local Similarity 46.2%; Pred. No. 5.4e-84;
Matches 238; Conservative 82; Mismatches 169; Indels 26; Gaps 7;

```
Qy      9 LLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLALA 63
      : | |  | |      : | |  | | | |  | |  | |  | :
Db      1 MAPALHWLLLWVGSGMLPAQGTHLGIRLPLRSLA-----GPPLGLRLPRETDEES----- 51

Qy     64 LEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPH 123
      |      : | : | | | | : | | : | | | | | | | | | | | | | |
Db     52 -----EEPGRRGSFVEMVDNLRGKSGQGYVEMTIGSPPQTLNILVDTGSSNFAVGAAPH 106

Qy    124 SYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTPKGFNTSFLVNIATIFES 183
      : :  | :  : | | | |  | | | | | | : | | | : | | :  | | | | |
Db    107 PFLHRYYYQRQLSSTYRDLRKGVYVPYTQGWEGELGTDLVSIHPGPNVTVRANIAAITES 166

Qy    184 ENFFLPGIKWNGILGLAYATLAKPSSSLETFFDLSLVQANIPNVFSMQMCGAGLPV---A 240
      : | | : |  | | | | | | : | : | | | | | | | : | | | | | :
Db    167 DKFFINGSNWEGLGLAYAEIARPDDSLPFFDLSLVKQTHIPNIFSLQLCGAGFPLNQTE 226

Qy    241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
      : | | : : : | | : | | | | : | | | : | : : | | | : | | : | |
Db    227 ALASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKS 286

Qy    301 IVDSGTTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISI 360
      | | | | | | | | | : | | : | | : : | | : | | | | | | | | | :
Db    287 IVDSGTTNLRPLPKVFEEAAVKSIIKAASSTEKFPDGFVLGEQLVCWQAGTTPWNIFPVISL 346

Qy    361 YLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGFYVIF 419
      | |  | : : : | | | | | | | | : : : : : : : | : | : | | | | : |
Db    347 YLMGEVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIMEGFYVVF 406

Qy    420 DRAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVC 479
      | | : | | : | | | | : : | | | | : : : | | : : : | : : : |
Db    407 DRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAAIC 466

Qy    480 GAILLV LIVLLLLPFRQCR--RPRDPEVVNDESSL 512
      | : : : | : : : | | | | : : : | | |
Db    467 -ALFMLPLCLMVCQWRCLRLRHQHDDFADDISLL 500
```

RESULT 11

Q9ULS1

ID Q9ULS1 PRELIMINARY; PRT; 532 AA.
AC Q9ULS1;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein KIAA1149 (Fragment).
GN KIAA1149.
OS Homo sapiens (Human).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20039618; PubMed=10574461;
RA Hirosawa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;
RT "Characterization of cDNA clones selected by the GeneMark analysis
RT from size-fractionated cDNA libraries from human brain.";
RL DNA Res. 6:329-336(1999).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR EMBL; AB032975; BAA86463.2; -.
DR HSSP; P56272; 1AM5.
DR GO; GO:0004194; F:pepsin A activity; IEA.
DR GO; GO:0008233; F:peptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR001969; Aspprotease_AS.
DR InterPro; IPR001461; Peptidase_A1.
DR InterPro; IPR009007; Pept_A_acid.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 1.
KW Hypothetical protein; Aspartyl protease; Hydrolase; Protease.
FT NON_TER 1 1
SQ SEQUENCE 532 AA; 58720 MW; 98B135D0D5FBD2E8 CRC64;

```

Qy	44	APTPGPGTPAERHADGLALAL-----EPALASPAGAAFLAMVDNLQGDSGRGYYLE	95
		: : : : : :	
Db	52	APSTASGCPCAAAWGGAPLGLRLPRETDEEP--EEPGRRGSFVEMVDNLRGKSGQGYVE	109
Qy	96	MLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWT	155
		: : : : : :	
Db	110	MTVGSPPTLNILVDTGSSNFAVGAAPHFPLHRYYQRQLSSTYRDLRKGVYPYTQGWKE	169
Qy	156	GFBVEDLVITPKGFNTSFLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFF	215
		: : : : : : :	
Db	170	GELGTDLVSIPHGPNVTVRANIAAITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFF	229
Qy	216	DSLVTQANIPNVFSMQMCGAGLPVAGS---GTNGGSLVLGGIEPSLYKGDIWYTPIKEEW	272
		: : : : : : : : : : : :	
Db	230	DSLWKQTHVPNLFSLQLCGAGFPLNQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREW	289
Qy	273	YYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVFDARVEAVARASLIPEF	332
		: : : : : : : : : : : : : :	
Db	290	YYEVIIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTNLRPLPKKVFEAAVKSIIKAASSTEKF	349
Qy	333	SDGFWTGSQACWTNSETPWSYFPKISIIYL RDENSSRSFRITILPOLYIQPMMGAGLNY-	391
		: : : : : : : : :	
Db	350	PDGFWLGEQLVCWQAGTTPWNIFPVISLYLMGEVTNQSFRTITILPQQYL RPVEDVATSQD	409
Qy	392	ECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCAETIAGA AVSEISGPFSTE	451
		: : : : : : : : : :	

```

Db      410 DCYKFAISQSSTGTVMGAVIMEGFYVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTL 469
Qy      452 DVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLIIVLLLLLPFRCQR--RPRDPEVVNDE 509
      |:      | | : : | : : : | : : : | : : | | | : : : |
Db      470 DMEDCGYNIPQTDESTLMTIAYVMAAIC-ALFMLPLCLMVCQWRCLRCLRQQHDDFADDI 528
Qy      510 SSL 512
      | |
Db      529 SLL 531

```

RESULT 12

Q8BQY4

```

ID      Q8BQY4      PRELIMINARY;      PRT;      501 AA.
AC      Q8BQY4;
DT      01-MAR-2003 (TrEMBLrel. 23, Created)
DT      01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT      01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE      Beta-site APP cleaving enzyme.
GN      BACE.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=C57BL/6J; TISSUE=Brain;
RX      MEDLINE=22354683; PubMed=12466851;
RA      The FANTOM Consortium,
RA      the RIKEN Genome Exploration Research Group Phase I & II Team;
RT      "Analysis of the mouse transcriptome based on functional annotation of
RT      60,770 full-length cDNAs.";
RL      Nature 420:563-573(2002).
DR      EMBL; AK046175; BAC32620.1; -.
DR      MGD; MGI:1346542; Bace.
DR      GO; GO:0004194; F:pepsin A activity; IEA.
DR      GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR      InterPro; IPR001969; Aspprotease_AS.
DR      InterPro; IPR001461; Peptidase_A1.
DR      InterPro; IPR009007; Pept_A_acid.
DR      Pfam; PF00026; asp; 1.
DR      PRINTS; PR00792; PEPSIN.
DR      PROSITE; PS00141; ASP_PROTEASE; 1.
SQ      SEQUENCE      501 AA;      55816 MW;      C0855513145E024E CRC64;

```

```

Query Match      44.0%;      Score 1183;      DB 11;      Length 501;
Best Local Similarity      46.0%;      Pred. No. 9.2e-84;
Matches 237;      Conservative      83;      Mismatches 169;      Indels      26;      Gaps      7;

```

```

Qy      9 LLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLALA 63
      : | |      | |      : | |      | | | |      |      | |      | | :
Db      1 MAPALHWLLLWVGSGMLPAQGTHLGIRLPLRSGLA-----GPPLGLRLPRETDEES----- 51
Qy      64 LEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPH 123
      |      : | : | | | | : | | : | | : | : | | | | | | | | | | | |
Db      52 -----EEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPQTLNILLVDTGSSNFAVGAAPH 106

```

Qy 124 SYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFES 183
 :: |: : |||| | | |||| | | :| |||:| | | : ||| | ||
 Db 107 PFLHRYYQRQLSSTYRDLRKGVYVPYTQGWEGELGTDLVSI PHGPNVTVRANIAAITES 166

Qy 184 ENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPV---A 240
 : ||: | | ||||| | :|:| ||| ||||| | :|||:| |:|:| ||| | :
 Db 167 DKFFINGSNWEGLGLAYAEIARPDDSLEPFFDSLVKQTHIPNIFSLQLCGAGFPLNQTE 226

Qy 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA 300
 : |||:::| |:| | :|||:| |||::: |:|:| | | :||:| ||| | :
 Db 227 ALASVGGSMIIGGIDHSlyTGRLWYTPIRREWYyEVIIVRVEINGQDLKMDCKEYNYDKS 286

Qy 301 IVDSGTTLLRLPQKVFDVAVEAVARASLIPEFSDFGTGSQLACWTNSETPWSYFPKISI 360
 ||||| | |||:| |:| | :| | | | | | | | | | | :| | | | :
 Db 287 IVDSGTTNLRLPKKVFEAAVKSikaASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISL 346

Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGFYVIF 419
 || | :|:| ||||| | :|:| : :|:| :| | : | :| :| ||||| :|
 Db 347 YLMGEVTNQSFRTILPQQYL RPVEDVATSQDDCYKFAVSQSSTGTVMGAVIMEGFYVVF 406

Qy 420 DRAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVC 479
 |||:| |:| | | : : ||| | | : | | :| : :| : :|
 Db 407 DRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAAIC 466

Qy 480 GAILLVLLIVLLLLPFCRQR--RPRDPEVVNDESSL 512
 | : : : | : : || | | : : : | | |
 Db 467 -ALFMLPLCLMVCQWRCLRLRHQHDDFADDISLL 500

RESULT 13

Q8IYC8

ID Q8IYC8 PRELIMINARY; PRT; 501 AA.
 AC Q8IYC8;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Beta-site APP-cleaving enzyme.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
 DR EMBL; BC036084; AAH36084.1; -.
 DR GO; GO:0004194; F:pepsin A activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR001461; Peptidase_A1.
 DR InterPro; IPR009007; Pept_A_acid.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 SQ SEQUENCE 501 AA; 55824 MW; 768595CF5517EFB7 CRC64;

Query Match 43.6%; Score 1172.5; DB 4; Length 501;
 Best Local Similarity 46.1%; Pred. No. 6.1e-83;
 Matches 239; Conservative 82; Mismatches 165; Indels 33; Gaps 9;

```

Qy      7 ALLLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLA 61
      |  || |  |:      : || |      |||      | |      ||
Db      2 AQALPWLLLWM---GAGVLPAGHTQHGI RLPLRSGLG-----GAPL-----GLR 42

Qy     62 LALE--PALASPAGAAANFLAMVDNLQGDSSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVA 119
      |  |      |      :|: ||||:| ||:|:|:| | :|:| | | ||||| |||||
Db     43 LPRETDEEPEEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPQTLNILVDTGSSNFAVG 102

Qy    120 GTPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVTI PKGFNTSFLVNIAT 179
      || :: |: : |||      | | ||| | | :| |||:| | | : |||
Db    103 AAPHPFLHRYYQRQLFSTYRDLRKG VYVPYTQGWEGELGTDLVSI PHGPNVTVRANIAA 162

Qy    180 IFESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPV 239
      | ||: ||: | | ||||| | :|:| ||| ||||| | :||:| |:|:| | | :
Db    163 ITESDKFFINGSNWEGILGLAYAEIARPD DSLEPFFDSL VKQTHVPNLFSLQLCGAGFPL 222

Qy    240 AGS---GTNGGSLVLGGIEPSLYKGD IWTPIKEEWYYQIEILKLEIGGQSLNLD CREYN 296
      |      : |||::| |: ||| | :||| | :|||:| | | :|:| | | :|:| |
Db    223 NQSEVLASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYN 282

Qy    297 ADKAIVDSGTTLLRLPQKVFD AVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFP 356
      ||:||||| |||:| |:| | :| | | | | | | | | | | | | | | | |
Db    283 YDKSIVDSGTTNLR LPKKVFEAAVKS IKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFP 342

Qy    357 KISYILRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGF 415
      ||:| | | :::| | | | | | | | :|:| : :|:| | | | :|:| :| | |
Db    343 VISLYLMGEVTNQSF RITILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGF 402

Qy    416 YVIFDRAQKRVGF AASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYAL 475
      ||:| | | :|:| | | | | : : ||| | | :      | | :|:| :
Db    403 YVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDESTLMTIAYVM 462

Qy    476 MSVCGAILLVLLIVLLLLLPFCRQR--RPRDPEVVNDESSL 512
      ::| |: :: : |: : :| | | | : : :| | |
Db    463 AAIC-ALFMLPLCLMVCQWRCLRLRQQHDDFADDISLL 500

```

RESULT 14

Q8C4F4

ID Q8C4F4 PRELIMINARY; PRT; 467 AA.
 AC Q8C4F4;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Beta-site APP cleaving enzyme.
 GN BACE.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum;

RX MEDLINE=22354683; PubMed=12466851;
 RA The FANTOM Consortium,
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs."
 RL Nature 420:563-573(2002).
 DR EMBL; AK082317; BAC38462.1; -.
 DR MGD; MGI:1346542; Bace.
 DR GO; GO:0004194; F:pepsin A activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR001461; Peptidase_A1.
 DR InterPro; IPR009007; Pept_A_acid.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 SQ SEQUENCE 467 AA; 52063 MW; 31AB674FF1843652 CRC64;

Query Match 39.0%; Score 1049; DB 11; Length 467;
 Best Local Similarity 41.7%; Pred. No. 2.4e-73;
 Matches 215; Conservative 77; Mismatches 163; Indels 60; Gaps 8;

Qy	9	LLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLALA	63
		: : :	
Db	1	MAPALHWLLLWVGSGMLPAQGTHLGIRLPLRSLA-----GPPLGLRLPRETDEES-----	51
Qy	64	LEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPH	123
		: :	
Db	52	-----EPPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPQTLNILVDTGSSNFAVGAAPH	106
Qy	124	SYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTPKGFNTSFLVNIATIFES	183
		: : : : : :	
Db	107	PFLHRYYYQRQLSSTYRDLRKGVYVPYTQGWEGELGTDLVSIHPGNVTVRANIAAITES	166
Qy	184	ENFFLPGIKWNGILGLAYATLAKPSSSLETFFDLSLVTQANI PN VF S M Q M C G A G L P V ---A	240
		: : : : : :	
Db	167	DKFFINGSNWE G I L G L A Y A E I A R P D D S L E P F F D S L V K Q T H I P N I F S L Q L C G A G F P L N Q T E	226
Qy	241	GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
		: : : : : : : :	
Db	227	ALASVGGSMIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKE-----	280
Qy	301	IVDSGTTLRLRPQKVFD A V V E A V A R A S L I P E F S D G F W T G S Q L A C W T N S E T P W S Y F P K I S I	360
		: :	
Db	281	-----TEKFPDGFWLGEQLVCWQAGTTPWNIFPVISL	312
Qy	361	YLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGFYVIF	419
		: : : : : : : : : :	
Db	313	YLMGEVTNQSFRTILPQQYL RP V E D V A T S Q D D C Y K F A V S Q S S T G T V M G A V I M E G F Y V V F	372
Qy	420	DRAQKRVGFAASPCAETIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVC	479
		: : : : : : : : :	
Db	373	DRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAAIC	432
Qy	480	GAILLV LIV L L L L P F R C Q R --RPRDPEVVNDESSL	512
		: : : : : : : : :	
Db	433	-ALFMLPLCLMVCQWRCLRLRHQHDDFADDISLL	466

RESULT 15

Q9CUU5

ID Q9CUU5 PRELIMINARY; PRT; 267 AA.
AC Q9CUU5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Adult male brain cDNA, RIKEN full-length enriched library,
DE clone:3526402A15 product:beta-site APP cleaving enzyme, full insert
DE sequence (Fragment).
GN BACE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=21085660; PubMed=11217851;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=99279253; PubMed=10349636;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning."
RL Meth. Enzymol. 303:19-44(1999).
RN [5]
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=20499374; PubMed=11042159;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Brain;
 RX MEDLINE=20530913; PubMed=11076861;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
 RA Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multicapillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 DR EMBL; AK014390; BAB29317.2; -.
 DR MGD; MGI:1346542; Bace.
 DR GO; GO:0004194; F:pepsin A activity; IEA.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
 DR InterPro; IPR001461; Peptidase_A1.
 DR InterPro; IPR009007; Pept_A_acid.
 DR Pfam; PF00026; asp; 1.
 FT NON_TER 1 1
 SQ SEQUENCE 267 AA; 30333 MW; 9413EB4530AB63B0 CRC64;

Query Match 24.3%; Score 653; DB 11; Length 267;
 Best Local Similarity 45.3%; Pred. No. 8.8e-43;
 Matches 121; Conservative 56; Mismatches 86; Indels 4; Gaps 3;

Qy 249 LVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKAIVDSGTTL 308
 ::|||: ||| | :||||: ||||: |::|| || | :||: || | :|||||
 Db 1 MIIGGIDHSLYTGSLWYTPIRREWYYEVIIVRVEINGQDLKMDCKEYNYDKSIVDSGTTN 60
 Qy 309 LRLPQKVFDAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISITL RDENSS 368
 ||||: |||: | |:: || :| |||| | || || |||: || ||: || | ::
 Db 61 LRLPKKVFEAAVKSIAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISLYLMGEVTN 120
 Qy 369 RSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVG 427
 :||||||| |::| : : ||: | : | : |::| :|||||: ||||: ||: |
 Db 121 QSFRITILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIMEGFYVVFDRARKRIG 180
 Qy 428 FAASPCAEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLI 487
 || | | : : ||| | | : | | : : | : : | : : :
 Db 181 FAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAAIC-ALFMLPL 239
 Qy 488 VLLLLPFRQQR--RPRDPEVVNDESSL 512
 |:: :|| | | : : : | | |
 Db 240 CLMVCQWRCLRCLRHQHDDFADDISLL 266

Search completed: March 4, 2004, 15:38:52
 Job time : 76.3936 secs

OM protein - protein search, using sw model

Run on: March 4, 2004, 15:22:30 ; Search time 16.5319 Seconds
(without alignments)
1631.532 Million cell updates/sec

Title: US-09-668-314C-2
Perfect score: 2687
Sequence: 1 MGALARALLPLLAQWLLRA.....RPRDPEVVNDESSLVRHRWK 518

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	2687	100.0	518	1	BAE2_HUMAN Q9y5z0 homo sapien
2	1187	44.2	501	1	BACE_RAT P56819 rattus norv
3	1185	44.1	501	1	BACE_MOUSE P56818 mus musculu
4	1178.5	43.9	501	1	BACE_HUMAN P56817 homo sapien
5	363.5	13.5	377	1	PEPC_MACFU P03955 macaca fusc
6	353	13.1	388	1	PEPC_HUMAN P20142 homo sapien
7	351.5	13.1	388	1	PEPC_CALJA Q9n2d3 callithrix
8	324.5	12.1	394	1	PEPC_CAVPO Q64411 cavia porce
9	320	11.9	402	1	RENI_MOUSE P06281 mus musculu
10	315.5	11.7	396	1	CATD_CLUHA Q9dex3 clupea hare
11	313.5	11.7	509	1	APR1_ORYSA Q42456 oryza sativ
12	313	11.6	392	1	PEPC_RAT P04073 rattus norv
13	310	11.5	383	1	PEPE_CHICK P16476 gallus gall
14	308.5	11.5	412	1	CATD_HUMAN P07339 homo sapien
15	306.5	11.4	410	1	CATD_MOUSE P18242 mus musculu
16	305.5	11.4	401	1	RENS_MOUSE P00796 mus musculu
17	305	11.4	407	1	CATD_RAT P24268 rattus norv

18	302	11.2	324	1	PEP1_GADMO	P56272	gadus morhu
19	302	11.2	405	1	CARP_YEAST	P07267	saccharomyc
20	301.5	11.2	398	1	CATE_RAT	P16228	rattus norv
21	300.5	11.2	387	1	PEP2_RABIT	P27821	oryctolagus
22	300.5	11.2	397	1	CATE_MOUSE	P70269	mus musculu
23	299	11.1	398	1	CATD_CHICK	Q05744	gallus gall
24	298.5	11.1	387	1	PEP4_RABIT	P28713	oryctolagus
25	298.5	11.1	400	1	RENI_SHEEP	P52115	ovis aries
26	297	11.1	388	1	PEPA_HUMAN	P00790	homo sapien
27	294.5	11.0	388	1	PEP2_MACFU	P27677	macaca fusc
28	291	10.8	388	1	PEP4_MACFU	P27678	macaca fusc
29	291	10.8	402	1	RENI_RAT	P08424	rattus norv
30	291	10.8	406	1	RENI_HUMAN	P00797	homo sapien
31	291	10.8	406	1	RENI_PANTR	P60016	pan troglod
32	290.5	10.8	396	1	CATE_RABIT	P43159	oryctolagus
33	289	10.8	387	1	PEP3_RABIT	P27822	oryctolagus
34	289	10.8	388	1	PAG_HORSE	Q28389	equus cabal
35	288.5	10.7	390	1	CATD_BOVIN	P80209	bos taurus
36	288	10.7	387	1	PEP1_RABIT	P28712	oryctolagus
37	288	10.7	388	1	PEP1_MACFU	P03954	macaca fusc
38	287	10.7	367	1	PEPA_CHICK	P00793	gallus gall
39	287	10.7	391	1	CATE_CAVPO	P25796	cavia porce
40	287	10.7	396	1	CATE_HUMAN	P14091	homo sapien
41	286	10.6	388	1	PEPA_MACMU	P11489	macaca mula
42	285.5	10.6	387	1	PEPA_CALJA	Q9n2d4	callithrix
43	285	10.6	396	1	CARP_NEUCR	Q01294	neurospora
44	284.5	10.6	386	1	PEPA_PIG	P00791	sus scrofa
45	283	10.5	388	1	PEPF_RABIT	P27823	oryctolagus

ALIGNMENTS

RESULT 1

BAE2_HUMAN

ID BAE2_HUMAN STANDARD; PRT; 518 AA.
AC Q9Y5Z0; Q9UJT6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Beta secretase 2 precursor (EC 3.4.23.45) (Beta-site APP-cleaving
DE enzyme 2) (Aspartyl protease 1) (Asp 1) (ASP1) (Membrane-associated
DE aspartic protease 1) (Memapsin-1) (Down region aspartic protease).
GN BACE2 OR ASP21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057170; PubMed=10591213;
RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;
RT "Membrane-anchored aspartyl protease with Alzheimer's disease
RT beta-secretase activity."
RL Nature 402:533-537(1999).

RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Bone marrow;
 RA Xin H., Stephans J.C., Duan X., Harrowe G., Kim E., Grieshammer U.,
 RA Giese K.;
 RT "Identification of a novel aspartic-like protease differentially
 RT expressed in human breast cancer cell lines.";
 RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Accarino M.P., Fumagalli P., Ottolenghi S., Taramelli R.;
 RT "Cloning of a gene from chromosome 21 Down region encoding a potential
 RT transmembrane aspartyl protease.";
 RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Solans A., Estivill X., de la Luna S.;
 RT "Cloning of a novel mammalian aspartyl protease.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20120043; PubMed=10656250;
 RA Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
 RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
 RA Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;
 RT "Identification of a novel aspartic proteinase (Asp 2) as
 RT beta-secretase.";
 RL Mol. Cell. Neurosci. 14:419-427(1999).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20144060; PubMed=10677483;
 RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of
 RT beta-amyloid precursor protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
 RN [7]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20289799; PubMed=10830953;
 RA Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
 RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
 RA Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
 RA Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
 RA Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
 RA Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;
 RT "The DNA sequence of human chromosome 21.";
 RL Nature 405:311-319(2000).
 RN [8]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Skin;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [9]
 RP CHARACTERIZATION.
 RX MEDLINE=22088158; PubMed=12093293;
 RA Turner R.T. III, Loy J.A., Nguyen C., Devasamudram T., Ghosh A.K.,
 RA Koelsch G., Tang J.;
 RT "Specificity of memapsin 1 and its implications on the design of
 RT memapsin 2 (beta-secretase) inhibitor selectivity."
 RL Biochemistry 41:8742-8746(2002).
 CC -!- CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-
 CC Val-Asn-Leu-|-Asp-Ala-Glu-Phe in the Swedish variant of
 CC Alzheimer's amyloid precursor protein.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF200342; AAF17078.1; -.
 DR EMBL; AF117892; AAD45240.1; -.
 DR EMBL; AF050171; AAD45963.1; -.
 DR EMBL; AF178532; AAF29494.1; -.
 DR EMBL; AF204944; AAF26368.1; -.
 DR EMBL; AF200192; AAF13714.1; -.
 DR EMBL; AL163284; CAB90458.1; -.
 DR EMBL; AL163285; CAB90554.1; -.
 DR EMBL; BC014453; AAH14453.1; -.
 DR HSSP; P00797; 2REN.
 DR MEROPS; A01.041; -.
 DR Genew; HGNC:934; BACE2.
 DR MIM; 605668; -.
 DR GO; GO:0005624; C:membrane fraction; TAS.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; TAS.
 DR GO; GO:0006464; P:protein modification; TAS.

DR GO; GO:0009306; P:protein secretion; TAS.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
 KW Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 ? POTENTIAL.
 FT CHAIN ? 518 BETA SECRETASE 2.
 FT DOMAIN 21 473 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 474 494 POTENTIAL.
 FT DOMAIN 495 518 CYTOPLASMIC (POTENTIAL).
 FT ACT_SITE 110 110 BY SIMILARITY.
 FT ACT_SITE 303 303 BY SIMILARITY.
 FT CARBOHYD 170 170 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 36 36 A -> T (IN REF. 6).
 SQ SEQUENCE 518 AA; 56180 MW; 2E903150823760D3 CRC64;

Query Match 100.0%; Score 2687; DB 1; Length 518;
 Best Local Similarity 100.0%; Pred. No. 2e-187;
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Db	1	MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPGPGTPAERHADGL	60
Qy	61	ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Db	61	ALALEPALASPAGAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAG	120
Qy	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTGQSWTGFGEDLVTIPKGENTSFLVNIATI	180
Db	121	TPHSYIDTYFDTERSSTYRSKGFDTVKYTGQSWTGFGEDLVTIPKGENTSFLVNIATI	180
Qy	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFEDSLVTQANIPNVFSMQMCGAGLPVA	240
Db	181	FESENFFLPGIKWNGILGLAYATLAKPSSSLETFEDSLVTQANIPNVFSMQMCGAGLPVA	240
Qy	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Db	241	GSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKA	300
Qy	301	IVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Db	301	IVDSGTTLRLRPQKVFDVAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI	360
Qy	361	YLRDENSRSRFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Db	361	YLRDENSRSRFRITILPQLYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFD	420
Qy	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480
Db	421	RAQKRVGFAASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCG	480

QY 481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518
 |||||
 Db 481 AILLVLIVLLLLPFRRCQRRPRDPEVVNDESSLVRHRWK 518

RESULT 2

BACE_RAT

ID BACE_RAT STANDARD; PRT; 501 AA.
 AC P56819;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Beta-secretase precursor (EC 3.4.23.46) (Beta-site APP cleaving
 DE enzyme) (Beta-site amyloid precursor protein cleaving enzyme)
 DE (Aspartyl protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic
 DE protease 2) (Memapsin-2).
 GN BACE.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20002972; PubMed=10531052;
 RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
 RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
 RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
 RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
 RA Treanor J., Rogers G., Citron M.;
 RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
 RT the transmembrane aspartic protease BACE.";
 RL Science 286:735-741(1999).
 CC -!- FUNCTION: Responsible for the proteolytic processing of the
 CC amyloid precursor protein (APP). Cleaves at the amino terminus of
 CC the A-beta peptide sequence, between residues 671 and 672 of APP,
 CC leads to the generation and extracellular release of beta-cleaved
 CC soluble APP, and a corresponding cell-associated carboxy-terminal
 CC fragment which is later release by gamma-secretase (By
 CC similarity).
 CC -!- CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-
 CC Val-Asn-Leu-|-Asp-Ala-Glu-Phe in the Swedish variant of
 CC Alzheimer's amyloid precursor protein.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF190727; AAF04144.1; -.
 DR HSSP; P32329; 1YPS.
 DR MEROPS; A01.004; -.
 DR InterPro; IPR001969; Aspprotease_AS.


```

          :| | | | :| | :| | | | | | : : | | | | | | : | | : : | : :
Db      405 VFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAA 464

QY      478 VCGAILLVLLIVLLLLLPFRCQR--RPRDPEVVNDESSL 512
          :| | : : : | : : :| | | | : : :| | |
Db      465 IC-ALFMLPLCLMVCQWRCLRLRHQHDDFADDISLL 500

```

RESULT 3

BACE_MOUSE

```

ID      BACE_MOUSE      STANDARD;      PRT;      501 AA.
AC      P56818;
DT      30-MAY-2000 (Rel. 39, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Beta-secretase precursor (EC 3.4.23.46) (Beta-site APP cleaving
DE      enzyme) (Beta-site amyloid precursor protein cleaving enzyme)
DE      (Aspartyl protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic
DE      protease 2) (Memapsin-2).
GN      BACE.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20002972; PubMed=10531052;
RA      Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
RA      Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
RA      Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
RA      Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
RA      Treanor J., Rogers G., Citron M.;
RT      "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
RT      the transmembrane aspartic protease BACE.";
RL      Science 286:735-741(1999).
RN      [2]
RP      REVISIONS TO 6 AND 81-87.
RA      Bennett B.D., Vassar R., Citron M.;
RL      Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20057170; PubMed=10591213;
RA      Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA      Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA      Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;
RT      "Membrane-anchored aspartyl protease with Alzheimer's disease
RT      beta-secretase activity.";
RL      Nature 402:533-537(1999).
CC      -!- FUNCTION: Responsible for the proteolytic processing of the
CC      amyloid precursor protein (APP). Cleaves at the amino terminus of
CC      the A-beta peptide sequence, between residues 671 and 672 of APP,
CC      leads to the generation and extracellular release of beta-cleaved
CC      soluble APP, and a corresponding cell-associated carboxy-terminal
CC      fragment which is later release by gamma-secretase (By
CC      similarity).
CC      -!- CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-
CC      Val-Asn-Leu-|-Asp-Ala-Glu-Phe in the Swedish variant of

```

CC Alzheimer's amyloid precursor protein.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: Brain.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF190726; AAF04143.2; -.
 DR EMBL; AF200346; AAF17082.1; -.
 DR HSSP; P56272; 1AM5.
 DR MEROPS; A01.004; -.
 DR MGD; MGI:1346542; Bace.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
 KW Signal.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 45 POTENTIAL.
 FT CHAIN 46 501 BETA-SECRETASE.
 FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 458 478 POTENTIAL.
 FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
 FT ACT_SITE 93 93 BY SIMILARITY.
 FT ACT_SITE 289 289 BY SIMILARITY.
 FT DISULFID 216 420 BY SIMILARITY.
 FT DISULFID 278 443 BY SIMILARITY.
 FT DISULFID 330 380 BY SIMILARITY.
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 501 AA; 55747 MW; C085A013145E474E CRC64;

Query Match 44.1%; Score 1185; DB 1; Length 501;
 Best Local Similarity 46.0%; Pred. No. 1.7e-78;
 Matches 237; Conservative 83; Mismatches 169; Indels 26; Gaps 7;

Qy 9 LLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTPGPGTPAERHADGLALA 63
 : | | | | : | | | | | | | | :
 Db 1 MAPALHWLLLWVGSGMLPAQGTHLGIRLPLRSGLA-----GPPLGLRLPRETDEES----- 51
 Qy 64 LEPALASPAGAANFLAMVDNLQGDSSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPH 123
 | : | : | | | : | | : | | : | : | | | | | | | | | |
 Db 52 -----EEPGRRGSFVEMVDNLRGKSGQGYVEMTVGSPPQTLNILVDTGSSNFAVGAAPH 106
 Qy 124 SYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLV TIPKGFNTSFLVNIATIFES 183
 : : | : : | | | | | | | | | | : | | | : | | : | | | | |

Db 107 PFLHRYYQRQLSSTYRDLRKG VYVPYTQGWEGELGTDLVSI PHGPNVTVRANIAAITES 166

Qy 184 ENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPV---A 240
: ||: | | ||||| :|:| ||| ||||| | :||:|:|:| ||| |:

Db 167 DKFFINGSNWEGILGLAYAEIARPDSDLFFDSL VKQTHIPNIFSLQLCGAGFPLNQTE 226

Qy 241 GSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYQIEILKLEIGGQSLNLD CREYNADKA 300
: |||:::| | | || | :|||: |||: |::| | | :||:| | |:

Db 227 ALASVGGSMIIGGIDHSLYTGSLWYTPIRREWYVEVIIVRVEINGQDLKMDCKEYNYDKS 286

Qy 301 IVDSGTTLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKISI 360
||||| |||:| | | :|:| || | | | |||: | | |:

Db 287 IVDSGTTNLR LPKKVFEAAVKS IKAASSTEKFPDGFWLGEQLVCWQAGTTPWNIFPVISL 346

Qy 361 YLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGFYVIF 419
|| | :||| ||| | | :|:| : ||:| :| | : ||| :||| |:

Db 347 YLMGEVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAVSQSSTGTVMGAVIMEGFYVVF 406

Qy 420 DRAQKRVGFAASPCAETIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVC 479
|||:|:| | | : : ||| | | : | | :|:| :|:

Db 407 DRARKRIGFAVSACHVHDEFRTAAVEGPFVTADMEDCGYNIPQTDESTLMTIAYVMAAIC 466

Qy 480 GAILLVLIIVLLLLPFR CQR--RPRDPEVVNDESSL 512
|: : : |:: :| | | : : : | | |

Db 467 -ALEMLPLCLMVCQWRCLRLRHQHDDFADDISLL 500

RESULT 4

BACE_HUMAN

ID BACE_HUMAN STANDARD; PRT; 501 AA.

AC P56817; Q9BYB9; Q9BYC0; Q9BYC1; Q9UJT5;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Beta-secretase precursor (EC 3.4.23.46) (Beta-site APP cleaving enzyme) (Beta-site amyloid precursor protein cleaving enzyme)

DE (Aspartyl protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2) (Memapsin-2).

GN BACE OR BACE1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A. (ISOFORM A).

RC TISSUE=Brain;

RX MEDLINE=20002972; PubMed=10531052;

RA Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,

RA Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,

RA Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,

RA Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,

RA Treanor J., Rogers G., Citron M.;

RT "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE.";

RL Science 286:735-741(1999).

RN [2]

RP SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 46-68, AND

RP CHARACTERIZATION.
 RC TISSUE=Brain;
 RX MEDLINE=20057171; PubMed=10591214;
 RA Sinha S., Anderson J.P., Barbour R., Basi G.S., Caccavello R.,
 RA Davis D., Doan M., Dovey H.F., Frigon N., Hong J., Jacobson-Croak K.,
 RA Jewett N., Keim P., Knops J., Lieberburg I., Power M., Tan H.,
 RA Tatsuno G., Tung J., Schenk D., Seubert P., Suomensaaari S.M., Wang S.,
 RA Walker D., Zhao J., McConlogue L., Varghese J.;
 RT "Purification and cloning of amyloid precursor protein beta-secretase
 RT from human brain.";
 RL Nature 402:537-540(1999).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM A).
 RX MEDLINE=20057170; PubMed=10591213;
 RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
 RA Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
 RA Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;
 RT "Membrane-anchored aspartyl protease with Alzheimer's disease beta-
 RT secretase activity.";
 RL Nature 402:533-537(1999).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM A).
 RX MEDLINE=20120043; PubMed=10656250;
 RA Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
 RA Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
 RA Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;
 RT "Identification of a novel aspartic proteinase (Asp 2) as beta-
 RT secretase.";
 RL Mol. Cell. Neurosci. 14:419-427(1999).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM B).
 RC TISSUE=Brain, and Pancreas;
 RA Michel B., De Pietri Tonelli D., Zacchetti D., Keller P.;
 RT "New beta-site APP cleaving enzyme isoform (BACE-1B) obtained from
 RT human brain and pancreas.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORM C).
 RC TISSUE=Pancreas;
 RA Zacchetti D., De Pietri Tonelli D., Schnurbus R.;
 RT "New beta-site APP cleaving enzyme isoform (BACE-1C) obtained from
 RT human pancreas.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORMS B; C AND D).
 RC TISSUE=Brain;
 RX MEDLINE=21408467; PubMed=11516562;
 RA Tanahashi H., Tabira T.;
 RT "Three novel alternatively spliced isoforms of the human beta-site
 RT amyloid precursor protein cleaving enzyme (BACE) and their effect on
 RT amyloid beta-peptide production.";
 RL Neurosci. Lett. 307:9-12(2001).
 RN [8]
 RP SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.
 RX MEDLINE=20144060; PubMed=10677483;
 RA Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
 RT "Human aspartic protease memapsin 2 cleaves the beta-secretase site of

RT beta-amyloid precursor protein.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
 RN [9]
 RP DISULFIDE BONDS.
 RX MEDLINE=21950860; PubMed=11953458;
 RA Fischer F., Molinari M., Bodendorf U., Paganetti P.;
 RT "The disulphide bonds in the catalytic domain of BACE are critical but
 RT not essential for amyloid precursor protein processing activity.";
 RL J. Neurochem. 80:1079-1088(2002).
 CC -!- FUNCTION: Responsible for the proteolytic processing of the
 CC amyloid precursor protein (APP). Cleaves at the amino terminus of
 CC the A-beta peptide sequence, between residues 671 and 672 of APP,
 CC leads to the generation and extracellular release of beta-cleaved
 CC soluble APP, and a corresponding cell-associated carboxy-terminal
 CC fragment which is later release by gamma-secretase.
 CC -!- CATALYTIC ACTIVITY: Broad endopeptidase specificity. Cleaves Glu-
 CC Val-Asn-Leu-|-Asp-Ala-Glu-Phe in the Swedish variant of
 CC Alzheimer's amyloid precursor protein.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=A; Synonyms=BACE-1A, BAC-501;
 CC IsoId=P56817-1; Sequence=Displayed;
 CC Name=B; Synonyms=BACE-1B, BACE-I-476;
 CC IsoId=P56817-2; Sequence=VSP_005223;
 CC Name=C; Synonyms=BACE-1C, BACE-I-457;
 CC IsoId=P56817-3; Sequence=VSP_005222;
 CC Name=D; Synonyms=BACE-1D, BACE-I-432;
 CC IsoId=P56817-4; Sequence=VSP_005222, VSP_005223;
 CC -!- TISSUE SPECIFICITY: Brain.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF190725; AAF04142.1; -.
 DR EMBL; AF201468; AAF18982.1; -.
 DR EMBL; AF200343; AAF17079.1; -.
 DR EMBL; AF204943; AAF26367.1; -.
 DR EMBL; AF338816; AAK38374.1; -.
 DR EMBL; AF338817; AAK38375.1; -.
 DR EMBL; AB050436; BAB40931.1; -.
 DR EMBL; AB050437; BAB40932.1; -.
 DR EMBL; AB050438; BAB40933.1; -.
 DR EMBL; AF200193; AAF13715.1; -.
 DR PIR; A59090; A59090.
 DR PDB; 1M4H; 28-AUG-02.
 DR MEROPS; A01.004; -.
 DR Genew; HGNC:933; BACE.
 DR MIM; 604252; -.
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.
 DR GO; GO:0008798; F:beta-aspartyl-peptidase activity; TAS..

DR GO; GO:0009405; P:pathogenesis; TAS.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 1.
 KW Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
 KW Signal; Alternative splicing; 3D-structure.
 FT SIGNAL 1 21 POTENTIAL.
 FT PROPEP 22 45
 FT CHAIN 46 501 BETA-SECRETASE.
 FT DOMAIN 22 457 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 458 478 POTENTIAL.
 FT DOMAIN 479 501 CYTOPLASMIC (POTENTIAL).
 FT ACT_SITE 93 93 BY SIMILARITY.
 FT ACT_SITE 289 289 BY SIMILARITY.
 FT DISULFID 216 420
 FT DISULFID 278 443
 FT DISULFID 330 380
 FT CARBOHYD 153 153 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 146 189 Missing (in isoform C and isoform D).
 FT /FTId=VSP_005222.
 FT VARSPLIC 190 214 Missing (in isoform B and isoform D).
 FT /FTId=VSP_005223.
 SQ SEQUENCE 501 AA; 55763 MW; 377CE4C824ACEF05 CRC64;

Query Match 43.9%; Score 1178.5; DB 1; Length 501;
 Best Local Similarity 46.2%; Pred. No. 5.1e-78;
 Matches 240; Conservative 82; Mismatches 164; Indels 33; Gaps 9;

Qy 7 ALLLPLLAQWLLRAAPELAPAPFT-----LPLRVAAATNRVVAPTGPGRTPAERHADGLA 61
 | || | |: : || | ||| | | ||
 Db 2 AQALPWLLLWM---GAGVLPAGHTQHGI RLP LRSGLG-----GAPL-----GLR 42

 Qy 62 LALE--PALASPAGAAANFLAMVDNLQGD SGRGYYLEMLIGTPPQKLQILVDTGSSNFAVA 119
 | | : |: ||||:| ||:| |:| | ||||| |||||
 Db 43 LPRETDEEPEEPGRRG SFVEMVDNLRGKSGQGYVEMTVGSPPQTLN ILVDTGSSNFAVG 102

 Qy 120 GTPHSYIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLV TIPKGFNTSFLVNIAT 179
 || :: |: : |||| | | |||| | | :| |||:| | | : |||
 Db 103 AAPHPFLHRY YQRQLSSTYRDLRKG VYVPYTQGWEGELGTDLV SIPHGPNVTVRANIAA 162

 Qy 180 IFESENFFLP GIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPV 239
 | ||: ||: | | ||||| |:| ||| ||||| | ::||:| |:| ||| |:
 Db 163 ITESDKFFINGSNWEGILGLAYAEIARPDDSLEPFFDSL VKQTHV PNLFSLQLCGAGFPL 222

 Qy 240 AGS---GTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYN 296
 | : |||::|||: ||| | :|||: ||||: |::|| || | :||:| |
 Db 223 NQSEVLASVGGSMIIGGIDHS LYTGSLWYTPIRREWYYEVII VRVEINGQDLKMDCKEYN 282

 Qy 297 ADKAIVDSGTTLLRLPQKVFD AVVEAVARASLIPEFS DGFWTGSQLACWTNSETPWSYFP 356
 ||:||||| ||||:| |:| |:: || :| |||| | || || |||: ||

Db 283 YDKSIVDSGTTNLRPLPKKVFEAAVKSIIKAASSTEKFPDGFVLGEQLVCWQAGTTPWNIIFP 342

Qy 357 KISIIYLRDENSSRSFRITILPQLYIQPMMGAGLNY-ECYRFGISPSTNALVIGATVMEGF 415
 ||:|| | ::||| |||| |::|: : :||:| || |: |:|| :|||

Db 343 VISLYLMGEVTNQSFRTILPQQYLRPVEDVATSQDDCYKFAISQSSTGTVMGAVIMEGF 402

Qy 416 YVIFDRAQKRVGFASPCAIEIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYAL 475
 ||:||||:||:|||| | | : : ||| | |: | | ::| :

Db 403 YVVFDRARKRIGFAVSACHVHDEFRTAAVEGPFVTLDMEDCGYNIPQTDESTLMTIAYVM 462

Qy 476 MSVCGAILLVLLIVLLLLPFCRQR--RPRDPEVVNDESSL 512
 ::| |: :: : |:: :|| | | : : :| | |

Db 463 AAIC-ALFMLPLCLMVCQWRCLRLRQHQHDDFADDISLL 500

RESULT 5

PEPC_MACFU

ID PEPC_MACFU STANDARD; PRT; 377 AA.

AC P03955;

DT 23-OCT-1986 (Rel. 02, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C) (Fragment).

GN PGC.

OS Macaca fuscata fuscata (Japanese macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;

OC Cercopithecinae; Macaca.

OX NCBI_TaxID=9543;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Gastric mucosa;

RX MEDLINE=92037645; PubMed=1935977;

RA Kageyama T., Tanabe K., Koiwai O.;

RT "Development-dependent expression of isozymogens of monkey

RT pepsinogens and structural differences between them.";

RL Eur. J. Biochem. 202:205-215(1991).

RN [2]

RP SEQUENCE OF 6-377.

RX MEDLINE=86168133; PubMed=3514597;

RA Kageyama T., Takahashi K.;

RT "The complete amino acid sequence of monkey progastricsin.";

RL J. Biol. Chem. 261:4406-4419(1986).

RN [3]

RP SEQUENCE OF 6-65.

RX MEDLINE=85289106; PubMed=3928607;

RA Kageyama T., Takahashi K.;

RT "Monkey pepsinogens and pepsins. VII. Analysis of the activation

RT process and determination of the NH2-terminal 60-residue sequence of

RT Japanese monkey progastricsin, and molecular evolution of

RT pepsinogens.";

RL J. Biochem. 97:1235-1246(1985).

CC -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but

CC shows preferential cleavage at Tyr-|-Xaa bonds; high activity

CC towards hemoglobin as substrate.

CC -!- PTM: Each pepsinogen is converted to corresponding pepsin at pH

CC 2.0 in part as a result of the release of a 47 aa activation

Db 193 ALTSPIFSVYLSAQ-----QGSSGGAVVFGGVDSSLYTGQIYWAPVTQELYWQIGIEEF 246

Qy 282 EIGGQSLNLD CREYNADKAI VDSGT TLLRLPQKVFD AVEAVARASLIPEFSDGFWTGSQ 341
 |||: | | :|||:|:| | :|: | :::| ||:|

Db 247 LIGGQASGW-CSE--GCQAIVDTGTSLLTVPQQYMSALLQA-----TGAQ 288

Qy 342 LACWTNSETPWSYF-----PKISIIYLRDENS SRSFRTILPQLYIQPMMGAGLNY 391
 | : | | :: : : | || ||

Db 289 -----EDEYGQFLVNCNSIQNLPTLTFII-----NGVEFPLPPSSYI-----LNN 328

Qy 392 ECY-RFGISP-----STNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
 | | : | | :::| : :| :::| : |||| :

Db 329 NGYCTVGVEPTYLSAQNSQPLWILGDVFLRSYYSVYDLSNNRVGFATA 376

RESULT 6

PEPC_HUMAN

ID PEPC_HUMAN STANDARD; PRT; 388 AA.

AC P20142;

DT 01-FEB-1991 (Rel. 17, Created)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C).

GN PGC.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88087276; PubMed=3335549;

RA Hayano T., Sogawa K., Ichihara Y., Fujii-Kuriyama Y., Takahashi K.;

RT "Primary structure of human pepsinogen C gene.";

RL J. Biol. Chem. 263:1382-1385(1988).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=89079679; PubMed=2909526;

RA Taggart R.T., Cass L.G., Mohandas T.K., Derby P., Barr P.J., Pals G.,

RA Bell G.I.;

RT "Human pepsinogen C (progastricsin). Isolation of cDNA clones,

RT localization to chromosome 6, and sequence homology with pepsinogen

RT A.";

RL J. Biol. Chem. 264:375-379(1989).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;

RX MEDLINE=89290840; PubMed=2567697;

RA Pals G., Azuma T., Mohandas T.K., Bell G.I., Bacon J.,

RA Samloff I.M., Walz D.A., Barr P.J., Taggart R.T.;

RT "Human pepsinogen C (progastricsin) polymorphism: evidence for a

RT single locus located at 6p21.1-pter.";

RL Genomics 4:137-148(1989).

RN [4]

RP SEQUENCE FROM N.A.

RA Wong R.N.S., Tang J.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.

RN [5]

RP SEQUENCE OF 17-101.
 RX MEDLINE=90130402; PubMed=2515193;
 RA Athauda S.B.P., Tanji M., Kageyama T., Takahashi K.;
 RT "A comparative study on the NH2-terminal amino acid sequences and
 RT some other properties of six isozymic forms of human pepsinogens and
 RT pepsins.";
 RL J. Biochem. 106:920-927(1989).
 RN [6]
 RP SEQUENCE OF 17-64.
 RX MEDLINE=83079318; PubMed=6816595;
 RA Foltmann B., Jensen A.L.;
 RT "Human progastricsin. Analysis of intermediates during activation
 RT into gastricsin and determination of the amino acid sequence of the
 RT propart.";
 RL Eur. J. Biochem. 128:63-70(1982).
 RN [7]
 RP X-RAY CRYSTALLOGRAPHY (1.62 ANGSTROMS).
 RX MEDLINE=95230687; PubMed=7714902;
 RA Moore S.A., Sielecki A.R., Chernaiia M.M., Tarasova N.I., James M.N.G.;
 RT "Crystal and molecular structures of human progastricsin at 1.62-A
 RT resolution.";
 RL J. Mol. Biol. 247:466-485(1995).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.36 ANGSTROMS).
 RX MEDLINE=98069649; PubMed=9406551;
 RA Khan A.R., Cherney M.M., Tarasova N.I., James M.N.;
 RT "Structural characterization of activation 'intermediate 2' on the
 RT pathway to human gastricsin.";
 RL Nat. Struct. Biol. 4:1010-1015(1997).
 CC -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
 CC shows preferential cleavage at Tyr-|-Xaa bonds; high activity
 CC towards hemoglobin as substrate.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M18667; AAA60062.1; ALT_INIT.
 DR EMBL; M18659; AAA60062.1; JOINED.
 DR EMBL; M18660; AAA60062.1; JOINED.
 DR EMBL; M18661; AAA60062.1; JOINED.
 DR EMBL; M18662; AAA60062.1; JOINED.
 DR EMBL; M18663; AAA60062.1; JOINED.
 DR EMBL; M18664; AAA60062.1; JOINED.
 DR EMBL; M18665; AAA60062.1; JOINED.
 DR EMBL; M18666; AAA60062.1; JOINED.
 DR EMBL; M23077; AAA60063.1; -.
 DR EMBL; M23069; AAA60063.1; JOINED.
 DR EMBL; M23070; AAA60063.1; JOINED.
 DR EMBL; M23071; AAA60063.1; JOINED.
 DR EMBL; M23072; AAA60063.1; JOINED.
 DR EMBL; M23073; AAA60063.1; JOINED.

DR EMBL; M23074; AAA60063.1; JOINED.
 DR EMBL; M23075; AAA60063.1; JOINED.
 DR EMBL; J04443; AAA60074.1; -.
 DR EMBL; U75272; AAB18273.1; -.
 DR PIR; A29937; A29937.
 DR PDB; 1HTR; 26-JAN-95.
 DR PDB; 1AVF; 25-FEB-98.
 DR MEROPS; A01.003; -.
 DR Genew; HGNC:8890; PGC.
 DR MIM; 169740; -.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR GO; GO:0004190; F:aspartic-type endopeptidase activity; TAS.
 DR GO; GO:0007586; P:digestion; TAS.
 DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal;
 KW 3D-structure.

FT	SIGNAL	1	16	
FT	PROPEP	17	59	ACTIVATION PEPTIDE.
FT	CHAIN	60	388	GASTRICSIN.
FT	ACT_SITE	91	91	
FT	ACT_SITE	276	276	
FT	DISULFID	104	109	
FT	DISULFID	267	271	
FT	DISULFID	310	343	
FT	CONFLICT	40	41	GE -> ED (IN REF. 6).
FT	CONFLICT	52	52	W -> S (IN REF. 6).
FT	STRAND	19	25	
FT	HELIX	29	35	
FT	TURN	36	37	
FT	HELIX	39	43	
FT	TURN	44	45	
FT	HELIX	50	54	
FT	HELIX	65	68	
FT	TURN	69	70	
FT	STRAND	73	79	
FT	TURN	80	83	
FT	STRAND	84	91	
FT	TURN	92	93	
FT	STRAND	97	101	
FT	TURN	102	103	
FT	HELIX	107	110	
FT	TURN	111	111	
FT	STRAND	115	115	
FT	HELIX	117	119	
FT	TURN	121	122	
FT	STRAND	124	134	
FT	TURN	135	136	
FT	STRAND	137	150	
FT	TURN	151	152	
FT	STRAND	153	163	
FT	HELIX	169	173	

FT	STRAND	178	181
FT	TURN	190	191
FT	HELIX	195	201
FT	TURN	202	203
FT	STRAND	209	214
FT	STRAND	221	227
FT	HELIX	232	234
FT	STRAND	235	244
FT	STRAND	251	254
FT	STRAND	256	259
FT	TURN	260	261
FT	STRAND	262	263
FT	TURN	266	269
FT	STRAND	271	275
FT	TURN	277	278
FT	STRAND	282	285
FT	HELIX	286	288
FT	HELIX	289	296
FT	TURN	297	297
FT	STRAND	299	300
FT	TURN	302	303
FT	STRAND	306	308
FT	HELIX	310	315
FT	STRAND	319	323
FT	TURN	324	325
FT	STRAND	326	330
FT	HELIX	332	335
FT	STRAND	336	338
FT	STRAND	343	345
FT	STRAND	347	350
FT	TURN	355	356
FT	STRAND	360	363
FT	HELIX	365	368
FT	TURN	369	370
FT	STRAND	371	376
FT	TURN	377	380
FT	STRAND	381	388
SQ	SEQUENCE	388 AA;	42426 MW; F862DFDC1438BB92 CRC64;

Query Match 13.1%; Score 353; DB 1; Length 388;
 Best Local Similarity 29.1%; Pred. No. 2.7e-18;
 Matches 120; Conservative 65; Mismatches 120; Indels 108; Gaps 21;

Qy	52	PAERHADG-LALALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVD	110
		:: :: : : : : :	
Db	50	PAWKYRFGDLSVTYEP-----MAYMD-----AAYFGEISIGTPPQNFLVLFD	91
Qy	111	TGSSNFAV-----AGTPHSYIDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDL	162
		: : ::	
Db	92	TGSSNLWVPSVYCQSQACTSHS----RFPNSESSTYSTNGQTFSLQYGSGLTGFFGYDT	147
Qy	163	VTIPKGFNTSFLVNIATIFESENFFLPG-----IKWNGILGLAYATLAKPSSSLETFFDS	217
		: : : : : : : :	
Db	148	LTV-----QSIQVPNQEFGLSEN--EPGTNFVYAQFDGIMGLAYPALSVDEAT--TAMQG	198
Qy	218	LVTQANIPN-VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQI	276
		: : : : : : : : : : : :	

```

Db      199 MVQEGALTSPVFSVYLSNQ-----QGSSGGAVVFGGVDSSLYTGQIYWAPVTQELYWQI 252
Qy      277 EILKLEIGGQSLNLD CREYNADKAIVDSGTLLRLPQKVFD AVEAVARASLIPEFSDGF 336
      | : |||: | | :|||:| |:| :||: |::|
Db      253 GIEEFLIGGQASGW-CSE--GCQAIVDTGTSLLTVPQQYMSALLQA----- 295
Qy      337 WTGSQLACWTNSETPWSYF-----PKISIIYLRDENS SRSFRTILPQLYIQPMMG 386
      ||:| | : | | : : | : | |
Db      296 -TGAQ-----EDEYGGQFLVNCNSIQNLPSLTFII-----NGVEFPLPPSSYI----- 336
Qy      387 AGLNYECY-RFGISP----STNA---LVIGATVMEGFYVIFDRAQKRVGFAAS 431
      |: | |: | | :| : :| :| : ||| :
Db      337 --LSNNGYCTVGVEPTYLSSQNGQPLWLGDVFLRSYYSVYDLGNNRVGFATA 387

```

RESULT 7

PEPC_CALJA

```

ID      PEPC_CALJA      STANDARD;      PRT;      388 AA.
AC      Q9N2D3;
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C).
GN      PGC.
OS      Callithrix jacchus (Common marmoset).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;      0
OC      Callithrix.
OX      NCBI_TaxID=9483;
RN      [1]
RP      SEQUENCE FROM N.A., SEQUENCE OF 17-26, FUNCTION, AND ENZYME
RP      REGULATION.
RC      TISSUE=Gastric mucosa;
RX      MEDLINE=20250834; PubMed=10788784;
RA      Kageyama T.;
RT      "New World monkey pepsinogens A and C, and prochymosins. Purification,
RT      characterization of enzymatic properties, cDNA cloning, and molecular
RT      evolution.";
RL      J. Biochem. 127:761-770(2000).
CC      -!- FUNCTION: Hydrolyzes a variety of proteins.
CC      -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
CC      shows preferential cleavage at Tyr-Xaa bonds; high activity
CC      towards hemoglobin as substrate.
CC      -!- ENZYME REGULATION: Inhibited by pepstatin.
CC      -!- MISCELLANEOUS: The optimal pH is around 2.
CC      -!- SIMILARITY: Belongs to peptidase family A1.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AB038385; BAA90872.1; -.
DR      PIR; JC7246; JC7246.

```

DR HSSP; P20142; 1AVF.
 DR MEROPS; A01.003; -.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal.
 FT SIGNAL 1 16
 FT PROPEP 17 59 ACTIVATION PEPTIDE (BY SIMILARITY).
 FT CHAIN 60 388 GASTRICSIN.
 FT ACT_SITE 91 91 BY SIMILARITY.
 FT ACT_SITE 276 276 BY SIMILARITY.
 FT DISULFID 104 109 BY SIMILARITY.
 FT DISULFID 267 271 BY SIMILARITY.
 FT DISULFID 310 343 BY SIMILARITY.
 SQ SEQUENCE 388 AA; 42503 MW; 0BC48DBD1F7D2D8C CRC64;

Query Match 13.1%; Score 351.5; DB 1; Length 388;
 Best Local Similarity 30.1%; Pred. No. 3.5e-18;
 Matches 112; Conservative 56; Mismatches 115; Indels 89; Gaps 17;

Qy 92 YYLEMLIGTPPQKLQILVDTGSSNFAV-----AGTPHSYIDTYFDTERSSTYRSKGF 143
 | : | : | | | | : | | | | | | | | : | | | | |
 Db 73 YFGEISIGTPPQNFLVLFDTGSSNLWVPSVYCQSQACTSHS----RFNPSASSTYSSNGQ 128

 Qy 144 DVTVKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATIFESENF LFG-----IKWNGILG 198
 : : : | | | | | : | : | | | | : : : | : |
 Db 129 TFSLQYGSGLTGFFGYDTLTV-----QSIQVPNQEFGLSEN--EPGTNFVYAQFDGIMG 181

 Qy 199 LAYATLAKPSSSLETFFDSLVTQANIPN-VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPS 257
 | | | | : : : | : : : : | | : | : | : : | : |
 Db 182 LAYPALSMGGAT--TAMQGMLQEGALTSPVFSFYLSNQ-----QGSSGGAVIFGGVDSS 233

 Qy 258 LYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKAIVDSGTTLLRLPQKVFD 317
 | | | | : : : | : | | | : | | | : | | : | : | : | : |
 Db 234 LYTGQIYWAPVTQELYWQIGIEEFLIGGQASGW-CSE--GCQAIVDTGTSLLTVPQQYMS 290

 Qy 318 AVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYF-----PKISIIYLRDENS 367
 | : | | | : | | : | | : | | : : :
 Db 291 AFLEA-----TGAQ-----EDEYGQFLVNCDSIQNLPTLTFII----- 323

 Qy 368 SRSFRITILPQLYIQPMMGAGLNYECY-RFGISP-----STNALVIGATVMEGFYVIF 419
 : | | | | : | | : | | : : | : : | : |
 Db 324 -NGVEFPLPPSSYI-----LSNNGYCTVGVEPTYLSSQNSQPLWILGDVFLRSYYSVF 375

 Qy 420 DRAQKRVGFAAS 431
 | | | | :
 Db 376 DLGNNRVGFATA 387

RESULT 8

PEPC_CAVPO

ID PEPC_CAVPO STANDARD; PRT; 394 AA.

AC Q64411;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C).
 GN PGC.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92355614; PubMed=1644829;
 RA Kageyama T., Ichinose M., Tsukada S., Miki K., Kurokawa K., Koiwai O.,
 RA Tanji M., Yakabe E., Athauda S.B., Takahashi K.;
 RT "Gastric procathepsin E and progastricsin from guinea pig.
 RT Purification, molecular cloning of cDNAs, and characterization of
 RT enzymatic properties, with special reference to procathepsin E.";
 RL J. Biol. Chem. 267:16450-16459(1992).
 CC -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
 CC shows preferential cleavage at Tyr-Xaa bonds; high activity
 CC towards hemoglobin as substrate.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; M88652; AAA37053.1; -.
 DR PIR; B43356; B43356.
 DR HSSP; P20142; 1AVF.
 DR MEROPS; A01.003; -.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal.
 FT SIGNAL 1 16 POTENTIAL.
 FT PROPEP 17 65 ACTIVATION PEPTIDE.
 FT CHAIN 66 394 GASTRICSIN.
 FT ACT_SITE 97 97 BY SIMILARITY.
 FT ACT_SITE 283 283 BY SIMILARITY.
 FT DISULFID 110 115 BY SIMILARITY.
 FT DISULFID 273 277 BY SIMILARITY.
 FT DISULFID 316 349 BY SIMILARITY.
 SQ SEQUENCE 394 AA; 42995 MW; 114F08E105D49865 CRC64;

Query Match 12.1%; Score 324.5; DB 1; Length 394;
 Best Local Similarity 29.0%; Pred. No. 3.2e-16;
 Matches 107; Conservative 63; Mismatches 116; Indels 83; Gaps 18;

Qy 92 YYLEMLIGTPPQKLQILVDTGSSNF-----AVAGTPHSYIDTYFDTERSSTYRSKGF 143
 |: :: :|||| |:| ||||| ::| | | | : ||| :

Db 79 YFGQISLGTPPQSFQVLFDTGSSNLWVPSVYCSSLACTTH----TRFNPRDSSTYVATDQ 134

Qy 144 DVTVKYTQGSWTGFVGEDLVTI-----PK-GFNTSFLVNIATIFESENFFLPG-----IK 192
 :::| || || | | :|| || | | | :| || :
 Db 135 SFSLEYGTGSLTGVFGYDTMTIQDIQVPKQEFGLS-----ETE----PGSDFVYAE 181

Qy 193 WNGILGLAYATLAKPSSSLETFFDLSVTQANI-PNVFSMQMCGAGLPVAGS--GTNGGSL 249
 ::|||| | |:: :: | | : : : ::||: : || |:: | |
 Db 182 FDGILGLGYPGLSEGGAT--TAMQGLLREGALSQSLFSVYL-----GSQQGSDEGQL 231

Qy 250 VLGgiePSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDcreYNADKAIVDSGTLL 309
 :|||:: ||| |||::||: :| |:|| | | : | : |||:|:|
 Db 232 ILGGVDESlytGDIYWTPVTQELYWQIGIEGFLIDGSASGWCSR--GCQGIVDTGTSL 288

Qy 310 RLPQKVFDavVEAVARASLIPEFSdGFWTGSQlACWtNsetPWSYfPKISiYLRdENSSR 369
 :| :||: | : : | :| : : | | :
 Db 289 TVPSDYLSTLVQAIGAE--NEYGEYF-----VSCSSIQDLPTLTFVISGV----- 332

Qy 370 SFRITILPQLYIQP-----MMGAGLNYECYREGISpSTN--ALVIGATVMEGFYVIFDRA 422
 : | || | :| | :|| :| : :| :|| |
 Db 333 --EFPLSPsAYILSGENYCMVGLESTY-----VSPGGGEPVWILGDVFLRSYYSVYDLA 384

Qy 423 QKRVGFAAS 431
 ||||| :
 Db 385 NNRVGFATA 393

RESULT 9

RENI_MOUSE

ID RENI_MOUSE STANDARD; PRT; 402 AA.

AC P06281; P97911; Q62153; Q62154;

DT 01-JAN-1988 (Rel. 06, Created)

DT 01-JAN-1988 (Rel. 06, Last sequence update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Renin 1 precursor (EC 3.4.23.15) (Angiotensinogenase) (Kidney renin).

GN REN1 OR REN-1 OR REN.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c;

RX MEDLINE=84182525; PubMed=6370686;

RA Holm I., Ollo R., Panthier J.-J., Rougeon F.;

RT "Evolution of aspartyl proteases by gene duplication: the mouse renin

RT gene is organized in two homologous clusters of four exons.";

RL EMBO J. 3:557-562(1984).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BALB/c; TISSUE=Kidney;

RX MEDLINE=90067953; PubMed=2685761;

RA Kim W.S., Murakami K., Nakayama K.;

RT "Nucleotide sequence of a cDNA coding for mouse Ren1 preprorenin.";

RL Nucleic Acids Res. 17:9480-9480(1989).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=DBA/2, and C57BL/10;
 RX MEDLINE=90108722; PubMed=2691339;
 RA Burt D.W., Mullins L.J., George H., Smith G., Brooks J., Pioli D.,
 RA Brammar W.J.;
 RT "The nucleotide sequence of a mouse renin-encoding gene, Ren-1d, and
 RT its upstream region.";
 RL Gene 84:91-104(1989).
 RN [4]
 RP SEQUENCE OF 1-30 FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=84298161; PubMed=6089205;
 RA Panthier J.-J., Dreyfus M., Roux D.T.L., Rougeon F.;
 RT "Mouse kidney and submaxillary gland renin genes differ in their 5'
 RT putative regulatory sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5489-5493(1984).
 RN [5]
 RP SEQUENCE OF 1-31 FROM N.A.
 RX MEDLINE=85085936; PubMed=6392850;
 RA Field L.J., Philbrick W.M., Howles P.N., Dickinson D.P.,
 RA McGowan R.A., Gross K.W.;
 RT "Expression of tissue-specific Ren-1 and Ren-2 genes of mice:
 RT comparative analysis of 5'-proximal flanking regions.";
 RL Mol. Cell. Biol. 4:2321-2331(1984).
 RN [6]
 RP SEQUENCE OF 22-37 AND 72-80.
 RC STRAIN=C57BL/10ROS X C3H/HEROS; TISSUE=Kidney;
 RX MEDLINE=97182599; PubMed=9030738;
 RA Jones C.A., Petrovic N., Novak E.K., Swank R.T., Sigmund C.D.,
 RA Gross K.W.;
 RT "Biosynthesis of renin in mouse kidney tumor As4.1 cells.";
 RL Eur. J. Biochem. 243:181-190(1997).
 CC -!- FUNCTION: Renin is a highly specific endopeptidase, whose only
 CC known function is to generate angiotensin I from angiotensinogen
 CC in the plasma, initiating a cascade of reactions that produce an
 CC elevation of blood pressure and increased sodium retention by the
 CC kidney.
 CC -!- CATALYTIC ACTIVITY: Cleaves Leu-|- bond in angiotensinogen to
 CC generate angiotensin I.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- TISSUE SPECIFICITY: Kidney.
 CC -!- INDUCTION: Renal renin is synthesized by the juxtaglomerular cells
 CC of the kidney in response to decreased blood pressure and sodium
 CC concentration.
 CC -!- POLYMORPHISM: In inbred mouse strains, there are at least two
 CC alleles which can occur at the Ren1 locus: Ren-1D and Ren-1C.
 CC The sequence shown is that of Ren-1C.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X00810; CAA25391.1; -.

DR EMBL; X00811; CAA25391.1; JOINED.
 DR EMBL; X00812; CAA25391.1; JOINED.
 DR EMBL; X00813; CAA25391.1; JOINED.
 DR EMBL; X00814; CAA25391.1; JOINED.
 DR EMBL; X00815; CAA25391.1; JOINED.
 DR EMBL; X00816; CAA25391.1; JOINED.
 DR EMBL; X00850; CAA25391.1; JOINED.
 DR EMBL; X00851; CAA25391.1; JOINED.
 DR EMBL; X16642; CAA34636.1; -.
 DR EMBL; K02596; AAA40045.1; -.
 DR EMBL; M32352; AAA40043.1; -.
 DR EMBL; K02800; AAA40044.1; -.
 DR EMBL; M34190; AAA40042.1; -.
 DR PIR; A00989; REMSK.
 DR HSSP; P00796; 1SMR.
 DR MEROPS; A01.007; -.
 DR MGD; MGI:97898; Ren1.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Plasma; Glycoprotein; Zymogen;
 KW Signal.
 FT SIGNAL 1 21
 FT PROPEP 22 71 ACTIVATION PEPTIDE.
 FT CHAIN 72 402 RENIN 1.
 FT ACT_SITE 102 102 BY SIMILARITY.
 FT ACT_SITE 287 287 BY SIMILARITY.
 FT DISULFID 115 122 BY SIMILARITY.
 FT DISULFID 278 282 BY SIMILARITY.
 FT CARBOHYD 69 69 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 139 139 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 58 58 W -> R (in Ren-1D).
 FT VARIANT 68 68 T -> I (in Ren-1D).
 FT VARIANT 160 160 S -> V (in Ren-1D).
 FT VARIANT 315 315 E -> D (in Ren-1D).
 FT VARIANT 352 352 N -> Y (in Ren-1D).
 FT CONFLICT 6 23 MISSING (IN REF. 1).
 FT CONFLICT 24 24 T -> I (IN REF. 1).
 FT CONFLICT 163 163 V -> VSRV (IN REF. 1).
 SQ SEQUENCE 402 AA; 44342 MW; D42920B555E97A38 CRC64;

Query Match 11.9%; Score 320; DB 1; Length 402;
 Best Local Similarity 28.6%; Pred. No. 7e-16;
 Matches 126; Conservative 66; Mismatches 181; Indels 68; Gaps 21;

Qy 10 LPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPG-PGTPAERHADGLALALE--- 65
 :|| | || :| |:|| | | : | || | | : |
 Db 6 MPLWALLLL-----WSPCTFSLPTRTATFERIPLKKMPVSVREILEERGVDMTRLSAEWGV 60
 Qy 66 ----PA---LASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAV 118
 | : | || | :| | | || | : ||||| : : ||||:| |
 Db 61 FTKRPSLTNLTSPVVLTYNL-----NTQ-----YYGEIGIGTPPQTFKVI FDTGSANLWV 110

Qy 119 AGTPHSY-----IDTYFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIPKGFNTS 172
 | | | : :: | | : | | : | | : | :
 Db 111 PSTKCSRLLYLACGIHSLYESSDSSSYMENGSDFTIHYGSGRVKGFLSQDSVTV-GGITVT 169

Qy 173 FLVNIATIFESENFFLPGIKWNGILGLAYATLAKPSSSLETFFDSLVTQANI-PNVFSMQ 231
 | | | : | : | : | : | : | : | :
 Db 170 QTFGEVTELPLIPFML--AKFDGVLGMGFP--AQAVGGVTPVFDHILSQGVLKEEVFSVY 225

Qy 232 MCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD 291
 | | | : | : | : | : | : | : | :
 Db 226 Y-----NRGSHLLGGEVVLGGSDPQHYQGNFHYVSISKTDWQITMKGVSVG--SSTLL 277

Qy 292 CREYNADKAIVDSGTTLRLPQKVFDVAVVEAV-ARASLIPEFSDFGTGSQLACWTNSET 350
 | | | : | : | : | : | : | : | :
 Db 278 CEEGCA--VVVDTGSSSFISAPTSSLKLIMQALGAKEKRIEY-----VVNC---SQV 324

Qy 351 PWSYFPKISIIYLRDENSSRSFRITILPQLYIQPMMGAGL-NYECYRFGISPSTNAL-VIG 408
 | | | : | : | : | : | : | : | :
 Db 325 P--TLPDISFDL----GGRAYTLSSDYVLQYPNRRDKLCTLALHAMDIPPPTGPVWVLG 378

Qy 409 ATVMEGFYVIFDRAQKRVGFA 429
 | | : | | | | : | | |
 Db 379 ATFIRKFYTEFDRHNNRIGFA 399

RESULT 10

CATD_CLUHA

ID CATD_CLUHA STANDARD; PRT; 396 AA.
 AC Q9DEX3;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Cathepsin D precursor (EC 3.4.23.5).
 OS Clupea harengus (Atlantic herring).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
 OC Clupea.
 OX NCBI_TaxID=7950;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Nielsen L.B., Stougaard P., Andersen P.S., Pedersen L.H.;
 RT "Cloning and sequence determination of herring muscle cathepsin D."
 RL Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE OF 62-82.
 RC TISSUE=Skeletal muscle;
 RX MEDLINE=21165469; PubMed=11207447;
 RA Nielsen L.B., Nielsen H.H.;
 RT "Purification and characterization of cathepsin D from herring muscle
 RT (Clupea harengus).";
 RL Comp. Biochem. Physiol. 128B:351-363(2001).
 CC -!- FUNCTION: Cathepsin D is an acid protease active in intracellular
 CC protein breakdown.
 CC -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
 CC that of pepsin A. Does not cleave the 4-Gln-|-His-5 bond in B
 CC chain of insulin.
 CC -!- ENZYME REGULATION: Inhibited by pepstatin.

CC -!- SUBUNIT: Monomer.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- MISCELLANEOUS: The isoelectric point is 6.8. Has optimal activity
 CC at pH 2.5 with hemoglobin as the substrate and the optimal
 CC temperature is 37 degrees Celsius.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF312364; AAG27733.1; -.
 DR HSSP; P07339; 1LYB.
 DR MEROPS; A01.009; -.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Signal; Zymogen.
 FT SIGNAL 1 18 POTENTIAL.
 FT PROPEP 19 61 ACTIVATION PEPTIDE.
 FT CHAIN 62 396 CATHEPSIN D.
 FT ACT_SITE 94 94 BY SIMILARITY.
 FT ACT_SITE 281 281 BY SIMILARITY.
 FT DISULFID 107 114 BY SIMILARITY.
 FT DISULFID 272 276 BY SIMILARITY.
 FT DISULFID 315 352 BY SIMILARITY.
 FT CARBOHYD 131 131 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 396 AA; 43315 MW; D0375DC38567A31B CRC64;

Query Match 11.7%; Score 315.5; DB 1; Length 396;
 Best Local Similarity 27.1%; Pred. No. 1.5e-15;
 Matches 112; Conservative 65; Mismatches 141; Indels 95; Gaps 18;

Qy 50 GTPAERHADGLALALEPALASPAGAAFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILV 109
 || : :| | : | :| |:: || |: :||| | ::
 Db 47 GTNSLQHNQGFSSNAP---TPETLKNYM-----DAQYYGEIGLGTVPVQMFVVF 93
 Qy 110 DTGSSNFAVAGTPHSYIDT-----YFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLV 163
 ||||| : | :| : :||| | : :| || :|:: :|
 Db 94 DTGSSNLWLPSIHCSFTDIACLLHHKYNGAKSSTYVKNGETEFAIQYGSGLSGYLSQDSC 153
 Qy 164 TIPKGFNTSFLVNIATIFESENFPLPGI-----KWNIGILGLAYATLAKPSSSLETFFDSL 218
 || :| :| | ||: |::|||:| | : : || :
 Db 154 TI-----GDIVVEKQLF-GEAIKQPGVAFIAAKFDGILGMAYPRIS--VDGVPPVFDMM 204
 Qy 219 VTQANI-PNVFSMQMCGAGLPVAGSGTN-----GGSVLVGGIEPSLYKGDIWYTPIKEEW 272
 ::| : ||| : | ||| :| | || | : :
 Db 205 MSQKKVEQNVSFYLF-----NRNPDTEPGGELLGGTDPKYYTGDFENYVPVTRQA 254
 Qy 273 YYQIEILKLEIGGQSLNLDREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLI--- 329

```

      |:|| : : || | | | |:: :|||:||:|: | | :|: ||
Db      255 YWQIHMDGMSIGSQ-LTL-CKD--GCEAIVDTGTSLITGPPAEVRALQKAIGAIPLIQGE 310

Qy      330 -----PEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSSRSFRITILPQ 378
      | | | | : | | | | |:: : : : |
Db      311 YMIDCKKVPTLPTIS--FNVGGK----TYSLTGEQY-----VLKESQGGKTICLSGLMG 358

Qy      379 LYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
      | | | | : :| : :| :||| |||| |
Db      359 LEIPP-----PAGPLWILGDVFIGQYYTVFDRESNRVGFAGS 395

```

RESULT 11

APR1_ORYSA

```

ID   APR1_ORYSA      STANDARD;          PRT;   509 AA.
AC   Q42456;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Aspartic proteinase oryzasin 1 precursor (EC 3.4.23.-).
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare / Japonica; TISSUE=Seed;
RX   MEDLINE=96048031; PubMed=7556174;
RA   Asakura T., Watanabe H., Abe K., Arai S.;
RT   "Rice aspartic proteinase, oryzasin, expressed during seed ripening
RT   and germination, has a gene organization distinct from those of
RT   animal and microbial aspartic proteinases.";
RL   Eur. J. Biochem. 232:77-83(1995).
CC   -!- DEVELOPMENTAL STAGE: Seed ripening and germination.
CC   -!- SIMILARITY: Belongs to peptidase family A1.
CC   -----
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL; D32165; BAA06876.1; -.
DR   EMBL; D32144; BAA06875.1; -.
DR   PIR; S66516; S66516.
DR   HSSP; P42210; 1QDM.
DR   MEROPS; A01.020; -.
DR   Gramene; Q42456; -.
DR   InterPro; IPR001969; Aspprotease_AS.
DR   InterPro; IPR009007; Pept_A_acid.
DR   InterPro; IPR001461; Peptidase_A1.
DR   InterPro; IPR007856; SapB_1.
DR   InterPro; IPR008138; SapB_2.
DR   InterPro; IPR008140; SapB_sub.

```

DR InterPro; IPR008373; Saposin.
 DR InterPro; IPR008139; SaposinB.
 DR Pfam; PF00026; asp; 1.
 DR Pfam; PF05184; SapB_1; 1.
 DR Pfam; PF03489; SapB_2; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PRINTS; PR01797; SAPOSIN.
 DR ProDom; PD001732; SapB_sub; 1.
 DR SMART; SM00118; SAPB; 2.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Zymogen; Glycoprotein; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 67 POTENTIAL.
 FT CHAIN 68 509 ASPARTIC PROTEINASE ORYZASIN 1.
 FT DOMAIN 318 416 SPECIFIC TO PLANT ASPARTIC PROTEINASES
 FT (BY SIMILARITY).
 FT ACT_SITE 103 103 BY SIMILARITY.
 FT ACT_SITE 290 290 BY SIMILARITY.
 FT DISULFID 116 122 BY SIMILARITY.
 FT DISULFID 281 285 BY SIMILARITY.
 FT DISULFID 428 465 BY SIMILARITY.
 FT CARBOHYD 252 252 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 509 AA; 54145 MW; 182F5DADA4CBE358 CRC64;

Query Match 11.7%; Score 313.5; DB 1; Length 509;
 Best Local Similarity 23.0%; Pred. No. 2.9e-15;
 Matches 127; Conservative 75; Mismatches 179; Indels 171; Gaps 19;

Qy 3 ALARALLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGPGPATPAERHADGLAL 62
 ::| || :| | || |: | | :|| | | | | | | |
 Db 5 SVALVLLAAVLLQALLPASAEGLVRIALKKRPIDENSRVAARLSG-----EEGARRLGL 59
 Qy 63 ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSN----- 115
 :| | : :|: : : | :|: :||| | :| | | | |
 Db 60 RGANSLGGGGGEGDIVALKNYMNAQ----YFGEIGVGTPPQKFTVIFDTGSSNLWVPSAK 115
 Qy 116 --FAVAGTPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLV TIPKGENTSF 173
 |::| || : : :|||: | :| || || || ||:
 Db 116 CYFSIACFFHS----RYKSGQSSTYQKNGKPAAIQYGTGSIAGFFESEDSVTVGD----- 165
 Qy 174 LVNIATIFESENF-----LPGI-----KWNIGLGLAYATLAKPSSSLETFFDSLVTQANI 224
 : : : | ||: |::||| | : : :
 Db 166 -----LVVKDQEFIEATKEPGLTFMVAKFDGILGLGFQEISVGDA-----V 206
 Qy 225 PNVFSMQMCG-AGLPVAGSGTN-----GGSLVLGGIEPSLYKGDIWYTPIKEEWYYQI 276
 | : | | || | || :| ||::|| |||: | |: : : |:
 Db 207 PVWYKMVEQGLVSEPVFSFWFNRRHSDEGEGGEIVFGMDPSHYKGNHTYVPVSQKGYWQF 266
 Qy 277 EILKLEIGGQSLNLD CREYNADKAIVDSGTTLLRLPQKVFDVAVVEAVARASLIPE----- 331
 |: : |||: | : || |||: || | : : | : : :
 Db 267 EMGDVLIGGKTTGF-CA--SGCSAIADSGTSLLAGPTAIITEINEKIGATGVVSQECKTV 323
 Qy 332 -----FSDGF----- 336
 |:
 Db 324 VSQYGGQILDLLLAETQPSKICSQVGLCTFDGKHGVSAGIKSVVDDEAGESNGLQSGPMC 383

Qy 337 -----WTGSQLACWTNSETPWSY-----FPKISIIYLRD 364
 | :||| : :| | :||| :
 Db 384 NACEMAVVWMQNQLAQNKTQDLILNYINQLCDKLPSMGESSVDCGSLASMPEISFTIGA 443

Qy 365 ENSSRSFRITILPQLYIQPMGAGLNYECY----RFGISPSTNAL-VIGATVMEGFYVIF 419
 : : : | : || :| | :| | | :| | : : :|
 Db 444 K-----KFALKPEEYIL-KVGEGAAAQCISGFTAMDIPPPRGPLWILGDVFMGAYHTVF 496

Qy 420 DRAQKRVGFAAS 431
 | : ||||| |
 Db 497 DYGKMRVGFAS 508

RESULT 12

PEPC_RAT

ID PEPC_RAT STANDARD; PRT; 392 AA.
 AC P04073;
 DT 01-NOV-1986 (Rel. 03, Created)
 DT 01-NOV-1986 (Rel. 03, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C).
 GN PGC.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar;
 RX MEDLINE=89255508; PubMed=2722863;
 RA Ishihara T., Ichihara Y., Hayano T., Katsura I., Sogawa K.,
 RA Fujii-Kuriyama Y., Takahashi K.;
 RT "Primary structure and transcriptional regulation of rat pepsinogen C
 RT gene.";
 RL J. Biol. Chem. 264:10193-10199(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Wistar;
 RX MEDLINE=87054020; PubMed=3780741;
 RA Ichihara Y., Sogawa K., Morohashi K., Fujii-Kuriyama Y., Takahashi K.;
 RT "Nucleotide sequence of a nearly full-length cDNA coding for
 RT pepsinogen of rat gastric mucosa.";
 RL Eur. J. Biochem. 161:7-12(1986).
 RN [3]
 RP SEQUENCE OF 16-112.
 RC STRAIN=Wistar;
 RX MEDLINE=84257697; PubMed=6743670;
 RA Arai K.M., Muto N., Tani S., Akahane K.;
 RT "The N-terminal sequence of rat pepsinogen.";
 RL Biochim. Biophys. Acta 788:256-261(1984).
 CC -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
 CC shows preferential cleavage at Tyr-Xaa bonds; high activity
 CC towards hemoglobin as substrate.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; M25993; AAA41827.1; -.
 DR EMBL; M25985; AAA41827.1; JOINED.
 DR EMBL; M25986; AAA41827.1; JOINED.
 DR EMBL; M25987; AAA41827.1; JOINED.
 DR EMBL; M25988; AAA41827.1; JOINED.
 DR EMBL; M25989; AAA41827.1; JOINED.
 DR EMBL; M25990; AAA41827.1; JOINED.
 DR EMBL; M25991; AAA41827.1; JOINED.
 DR EMBL; M25992; AAA41827.1; JOINED.
 DR EMBL; X04644; CAA28305.1; -.
 DR PIR; A33510; A24608.
 DR HSSP; P20142; 1AVF.
 DR MEROPS; A01.003; -.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal.
 FT SIGNAL 1 16
 FT PROPEP 17 62 ACTIVATION PEPTIDE.
 FT CHAIN 63 392 GASTRICSIN.
 FT ACT_SITE 94 94
 FT ACT_SITE 280 280
 FT DISULFID 107 112 BY SIMILARITY.
 FT DISULFID 270 275 BY SIMILARITY.
 FT DISULFID 314 347 BY SIMILARITY.
 FT CONFLICT 31 31 E -> Q (IN REF. 3).
 FT CONFLICT 103 103 S -> A (IN REF. 3).
 FT CONFLICT 109 109 S -> L (IN REF. 3).
 SQ SEQUENCE 392 AA; 42833 MW; 092A5EAF2783EDD1 CRC64;

Query Match 11.6%; Score 313; DB 1; Length 392;
 Best Local Similarity 29.5%; Pred. No. 2.2e-15;
 Matches 105; Conservative 56; Mismatches 139; Indels 56; Gaps 16;

Qy 92 YYLEMLIGTPPQKLQILVDTGSSNFAV-----AGTPHSYIDTYFDTERSSTYRSKGF 143
 |: |: ||||| |: ||||| | | |: |: :||| ::|
 Db 76 YFGEISIGTPPQNFLVLFDTGSSNLWVSSVYCQSEACTTHA----RFNPSKSSTYYTEGQ 131
 Qy 144 DVTVKYTQGSWTGFVGEDLVTPKGFNTSFLVNIATIFESENFPLPG-----IKWNGILG 198
 :::| || ||| | | :|: | | || | | ::||:|
 Db 132 TFSLQYGTGSLTGFFGYDTLTV-----QSIQVPNQEFGLSEN--EPGTNFVYAQFDGIMG 184
 Qy 199 LAYATLAKPSSSLETFFDLSVTQANIPNVFSMQMCGAGLPVAGS--GTNGGSLVLGGIEP 256
 ||| |: | | : : | : | | | :||| :| ||::
 Db 185 LAYPGLS--SGGATTALQGMLGE----GALSQPLFGVYL--GSQQGSNGGQIVFEGGVDK 235
 Qy 257 SLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLDCREYNADKAIVDSGTTLLRLPQKVF 316
 :|| |:| : |: :| |:|| | || |: | : |||:|:|:| :| :

```

Db      236 NLYTGEITWVPVTQELYWQITIDDFLIGDQASGW-CSSQGC-QGIVDTGTSLLVMPAQYL 293
Qy      317 DAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISYLRDENSSRSFRITIL 376
      ::: : |: : | ::: | : | : | : :
Db      294 SELLQTIGAQE--GEYGEYF-----VSCDSVSS-----LPTLSFVL-----NGVQFPLS 335
Qy      377 PQLY-IQPMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAAS 431
      | | || : | ::: : : | || : || | |
Db      336 PSSYIIQEDNFCMVGLESLTSESGQPLWILGDVFLRSYYAIFDMGNNKVGLATS 391

```

RESULT 13

PEPE_CHICK

```

ID  PEPE_CHICK      STANDARD;      PRT;      383 AA.
AC  P16476;
DT  01-AUG-1990 (Rel. 15, Created)
DT  01-AUG-1990 (Rel. 15, Last sequence update)
DT  28-FEB-2003 (Rel. 41, Last annotation update)
DE  Embryonic pepsinogen precursor (EC 3.4.23.-).
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC  Gallus.
OX  NCBI_TaxID=9031;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=88227903; PubMed=3131317;
RA  Hayashi K., Agata K., Mochii M., Yasugi S., Eguchi G., Mizuno T.;
RT  "Molecular cloning and the nucleotide sequence of cDNA for embryonic
RT  chicken pepsinogen: phylogenetic relationship with prochymosin.";
RL  J. Biochem. 103:290-296(1988).
CC  -!- DEVELOPMENTAL STAGE: Specifically secreted during the embryonic
CC      period in the chicken proventriculus (glandular stomach).
CC  -!- SIMILARITY: Belongs to peptidase family A1.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; D00215; BAA00153.1; -.
DR  PIR; A41443; A41443.
DR  HSSP; P00794; 4CMS.
DR  MEROPS; A01.028; -.
DR  InterPro; IPR001969; Aspprotease_AS.
DR  InterPro; IPR009007; Pept_A_acid.
DR  InterPro; IPR001461; Peptidase_A1.
DR  Pfam; PF00026; asp; 1.
DR  PRINTS; PR00792; PEPSIN.
DR  PROSITE; PS00141; ASP_PROTEASE; 2.
KW  Hydrolase; Aspartyl protease; Digestion; Signal; Glycoprotein.
FT  SIGNAL      1      16      POTENTIAL.
FT  CHAIN      17      383      EMBRYONIC PEPSINOGEN.
FT  ACT_SITE    94      94      BY SIMILARITY.

```

FT ACT_SITE 276 276 BY SIMILARITY.
 FT DISULFID 107 112 BY SIMILARITY.
 FT DISULFID 267 271 BY SIMILARITY.
 FT DISULFID 310 344 BY SIMILARITY.
 FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 309 309 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 350 350 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 51 51 T -> S.
 SQ SEQUENCE 383 AA; 41719 MW; 1642796871611F54 CRC64;

Query Match 11.5%; Score 310; DB 1; Length 383;
 Best Local Similarity 26.8%; Pred. No. 3.5e-15;
 Matches 106; Conservative 63; Mismatches 136; Indels 90; Gaps 15;

Qy 56 HA--DGLALALEPALASPAGAAFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGS 113
 || | | : || | | | | | | | : ||||| :: ||||
 Db 55 HAFPDVLTVVTEPLL-----NTLDM-----EYYGTISIGTPPQDFTVVFDTS 97

 Qy 114 SNFAVAG----TPHSYIDTYFDTERSSTYRSKGFDTVVKYTQGSWTGFVGEDLVITPKGF 169
 || | : | | : ||||| : | | : : : | | | | | :
 Db 98 SNLWVPSVSCTSPACQSHQMFNPSQSSTYKSTGQNLSIHYGTGDMEGTVGCDTVTVASLM 157

 Qy 170 NTSFLVNIATIFESENFFLPGLKWNGLGLAYATLAKPSSSLETFDLSLTQANI-PNVF 228
 : | : | : | | : : | : | : | : | : | : | : | : | : | : | :
 Db 158 DTNQLFGLST-SEPGQFFV-YVKFDGILGLGYPSLA--ADGITPVFDNMVNESLLEQNLF 213

 Qy 229 SMQMCAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSL 288
 | : : : | : | | | : | : | : | : | : | : | : | :
 Db 214 SVYLSREPM-----GSMVVFEGGIDESYFTGSINWIPVSYQGYWQISMDSIIVNKQEI 265

 Qy 289 NLDCREYNADKAIVDSGTTLRLPQKVFDAVVEAVARASLIPEFSDGFWTGSQACWTNS 348
 : : | : | : | : | : | : | : | : | : | : | : | : | :
 Db 266 ACS----SGCQAIIDTGTSLVAGPASDINDIQSAVG-----ANQ 300

 Qy 349 ETPWSYFPKISIIYLRDENSSRSFRITILPQLYIQPMMGAGLNY-----ECY 394
 | | | | | : | : : : | : | : | : | : | : | : | :
 Db 301 NTYGEY-----SVNCSHILAMPDVVF--VIG-GIQYPVPALAYTEQNGQGTCTM 345

 Qy 395 RFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFA 429
 : | : : | : : | | | | | | | | | |
 Db 346 SSFQNSSADLWILGDVFIRVYYSIFDRANNRVGLA 380

RESULT 14

CATD_HUMAN

ID CATD_HUMAN STANDARD; PRT; 412 AA.

AC P07339;

DT 01-APR-1988 (Rel. 07, Created)

DT 01-APR-1988 (Rel. 07, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Cathepsin D precursor (EC 3.4.23.5).

GN CTSD.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=85270436; PubMed=3927292;
 RA Faust P.L., Kornfeld S., Chirgwin J.M.;
 RT "Cloning and sequence analysis of cDNA for human cathepsin D.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:4910-4914(1985).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=87231068; PubMed=3588310;
 RA Westley B.R., May F.E.B.;
 RT "Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive
 RT human breast cancer cells.";
 RL Nucleic Acids Res. 15:3773-3786(1987).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91299158; PubMed=2069717;
 RA Redecker B., Heckendorf B., Grosch H.W., Mersmann G., Hasilik A.;
 RT "Molecular organization of the human cathepsin D gene.";
 RL DNA Cell Biol. 10:423-431(1991).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP SEQUENCE OF 1-22 FROM N.A.
 RX MEDLINE=94085791; PubMed=8262386;
 RA May F.E., Smith D.J., Westley B.R.;
 RT "The human cathepsin D-encoding gene is transcribed from an estrogen-
 RT regulated and a constitutive start point.";
 RL Gene 134:277-282(1993).
 RN [6]
 RP SEQUENCE OF 1-22 FROM N.A.
 RX MEDLINE=95021301; PubMed=7935485;
 RA Augereau P., Miralles F., Cavailles V., Gaudalet C., Parker M.,
 RA Rochefort H.;
 RT "Characterization of the proximal estrogen-responsive element of
 RT human cathepsin D gene.";

RL Mol. Endocrinol. 8:693-703(1994).
 RN [7]
 RP SEQUENCE OF 170-180.
 RC TISSUE=Liver;
 RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
 RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
 RA Appel R.D., Hughes G.J.;
 RL Submitted (JUN-1992) to Swiss-Prot.
 RN [8]
 RP CARBOHYDRATE-LINKAGE SITE ASN-263.
 RX MEDLINE=22660472; PubMed=12754519;
 RA Zhang H., Li X.-J., Martin D.B., Aebersold R.;
 RT "Identification and quantification of N-linked glycoproteins using
 RT hydrazide chemistry, stable isotope labeling and mass spectrometry.";
 RL Nat. Biotechnol. 21:660-666(2003).
 RN [9]
 RP VARIANT VAL-58.
 RX MEDLINE=20179010; PubMed=10716266;
 RA Papassotiropoulos A., Bagli M., Kurz A., Kornhuber J., Forstl H.,
 RA Maier W., Pauls J., Lautenschlager N., Heun R.;
 RT "A genetic variation of cathepsin D is a major risk factor for
 RT Alzheimer's disease.";
 RL Ann. Neurol. 47:399-403(2000).
 RN [10]
 RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
 RC TISSUE=Spleen;
 RX MEDLINE=93223670; PubMed=8467789;
 RA Metcalf P., Fusek M.;
 RT "Two crystal structures for cathepsin D: the lysosomal targeting
 RT signal and active site.";
 RL EMBO J. 12:1293-1302(1993).
 RN [11]
 RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
 RC TISSUE=Liver;
 RX MEDLINE=93342076; PubMed=8393577;
 RA Baldwin E.T., Bhat T.N., Gulnik S., Hosur M.V., Sowder R.C. II,
 RA Cachau R.E., Collins J., Silva A.M., Erickson J.W.;
 RT "Crystal structures of native and inhibited forms of human cathepsin
 RT D: implications for lysosomal targeting and drug design.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6796-6800(1993).
 CC -!- FUNCTION: Acid protease active in intracellular protein breakdown.
 CC Involved in the pathogenesis of several diseases such as breast
 CC cancer and possibly Alzheimer's disease.
 CC -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
 CC that of pepsin A. Does not cleave the 4-Gln-|-His-5 bond in B
 CC chain of insulin.
 CC -!- SUBUNIT: Consists of a light chain and a heavy chain.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- POLYMORPHISM: The Val-58 allele is significantly overrepresented
 CC in demented patients (11.8%) compared with nondemented controls
 CC (4.9%). Carriers of the Val-58 allele have a 3.1-fold increased
 CC risk for developing AD than noncarriers.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M11233; AAB59529.1; -.
 DR EMBL; X05344; CAA28955.1; -.
 DR EMBL; M63138; AAA51922.1; -.
 DR EMBL; M63134; AAA51922.1; JOINED.
 DR EMBL; M63135; AAA51922.1; JOINED.
 DR EMBL; M63136; AAA51922.1; JOINED.
 DR EMBL; M63137; AAA51922.1; JOINED.
 DR EMBL; BC016320; AAH16320.1; -.
 DR EMBL; L12980; AAA16314.1; -.
 DR EMBL; S74689; AAD14156.1; -.
 DR EMBL; S52557; AAD13868.1; -.
 DR PIR; A25771; KHHUD.
 DR PDB; 1LYA; 31-JAN-94.
 DR PDB; 1LYB; 31-JAN-94.
 DR PDB; 1LYW; 22-JUL-99.
 DR MEROPS; A01.009; -.
 DR SWISS-2DPAGE; P07339; HUMAN.
 DR Siena-2DPAGE; P07339; -.
 DR Genew; HGNC:2529; CTSD.
 DR MIM; 116840; -.
 DR GO; GO:0004192; F:cathepsin D activity; TAS.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Signal; Zymogen;
 KW Polymorphism; Alzheimer's disease; 3D-structure.
 FT SIGNAL 1 18
 FT PROPEP 19 64 ACTIVATION PEPTIDE.
 FT CHAIN 65 412 CATHEPSIN D.
 FT CHAIN 65 161 CATHEPSIN D LIGHT CHAIN (PROBABLE).
 FT CHAIN 169 412 CATHEPSIN D HEAVY CHAIN (PROBABLE).
 FT ACT_SITE 97 97
 FT ACT_SITE 295 295
 FT DISULFID 91 160
 FT DISULFID 110 117
 FT DISULFID 286 290
 FT DISULFID 329 366
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .).
 FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .).
 FT VARIANT 58 58 A -> V (ASSOCIATED WITH INCREASED RISK IN
 FT AD; POSSIBLY INFLUENCES SECRETION AND
 FT INTRACELLULAR MATURATION; dbSNP:17571).
 FT /FTId=VAR_011621.
 FT STRAND 67 74
 FT TURN 75 77
 FT STRAND 78 85
 FT TURN 86 89
 FT STRAND 90 97
 FT TURN 98 99

FT	STRAND	103	107
FT	TURN	108	109
FT	TURN	112	113
FT	HELIX	115	118
FT	TURN	119	119
FT	STRAND	123	123
FT	HELIX	125	127
FT	TURN	129	130
FT	STRAND	132	141
FT	STRAND	146	158
FT	STRAND	172	184
FT	HELIX	188	192
FT	STRAND	197	200
FT	HELIX	204	206
FT	HELIX	208	210
FT	HELIX	214	220
FT	TURN	221	222
FT	STRAND	228	233

Query Match 11.5%; Score 308.5; DB 1; Length 412;
 Best Local Similarity 27.1%; Pred. No. 5e-15;
 Matches 121; Conservative 75; Mismatches 180; Indels 71; Gaps 22;

Qy	9	LLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTPG-----PGTPAERHADGLAL	62
		: : :: :: :	
Db	6	LLPLAL--CLLAAP--ASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPVSKYSQAVPA	61
Qy	63	ALEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTP	122
		: :: : ::	
Db	62	VTEGPI--PEVLKNYM-----DAQYYGEIGIGTPPQCFTVVFDTGSSNLWVPSIH	109
Qy	123	HSYIDT-----YFDTERSSTYRSKGFDTVKYTQGSWTGFVGEDLVTIP--KGFNTSFL	174
		: :::: : : :: : :: :	
Db	110	CKLLDIACWIIHKYNSDKSSTYVKNGTSEFDIHYGSGSLSGYLSQDTVSVPCQSASSASAL	169
Qy	175	--VNIATIFESENFFLPGI-----KWNIGILGLAYATLAKPSSSLETFFDSLVTQANI-PN	226
		: :: :: :: : : :	
Db	170	GGVKVERQVFGKATKQPGITFIAAKFDGILGMAYPRIS--VNNVLPVFDNLMQQKLVDQN	227
Qy	227	VFSMQMCGAGLPVAGSGTNGGSLVLGGIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQ	286
		: : : : : : : : :: :	
Db	228	IFSFY-----LSRDPDAQPGGELMLGGTDSKYYKGSLSYLVNTRKAYWQVHLDQVEV-AS	281
Qy	287	SLNLDCREYNADKAIVDSGTLLRLPQKVFDVAVVEAVARASLIPEFSDGFWTGSQACWT	346
		: : : : : : :	
Db	282	GLTL-CKE--GCEAIVDTGTSLMVGP--VDEVRELQKAIGAVPLIQGEY----MIPC--	329
Qy	347	NSETPWSYFPKISIIYLRDENSSRSFRITILPQLYIQPMMGAGLNYECYRF---GISPSTN	403
		:: : :::: : : :	
Db	330	---EKVSTLPAITLKL----GGKGYKLS--PEDYTLKVSQAGKTLCLSGFMGMDIPPPSG	380
Qy	404	AL-VIGATVMEGFYVIFDRAQKRVGFA	429
		:: : : :	
Db	381	PLWILGDVFIGRYYTVFDRDNNRVGFA	407

CATD_MOUSE

ID CATD_MOUSE STANDARD; PRT; 410 AA.
AC P18242;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
GN CTSD.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=91088345; PubMed=2263503;
RA Diedrich J.F., Staskus K.A., Retzel E.F., Haase A.T.;
RT "Nucleotide sequence of a cDNA encoding mouse cathepsin D.";
RL Nucleic Acids Res. 18:7184-7184(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90326544; PubMed=2374732;
RA Grusby M.J., Mitchell S.C., Glimcher L.H.;
RT "Molecular cloning of mouse cathepsin D.";
RL Nucleic Acids Res. 18:4008-4008(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RX MEDLINE=94280622; PubMed=8011168;
RA Hetman M., Perschl A., Saftig P., von Figura K., Peters C.;
RT "Mouse cathepsin D gene: molecular organization, characterization of
RT the promoter, and chromosomal localization.";
RL DNA Cell Biol. 13:419-427(1994).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain, and Mammary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- FUNCTION: Acid protease active in intracellular protein breakdown.
 CC -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
 CC that of pepsin A. Does not cleave the 4-Gln-|-His-5 bond in B
 CC chain of insulin.
 CC -!- SUBUNIT: Consists of a light chain and a heavy chain.
 CC -!- SUBCELLULAR LOCATION: Lysosomal.
 CC -!- SIMILARITY: Belongs to peptidase family A1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X53337; CAA37423.1; -.
 DR EMBL; X52886; CAA37067.1; -.
 DR EMBL; X68378; CAA48453.1; -.
 DR EMBL; X68379; CAA48453.1; JOINED.
 DR EMBL; X68380; CAA48453.1; JOINED.
 DR EMBL; X68381; CAA48453.1; JOINED.
 DR EMBL; X68382; CAA48453.1; JOINED.
 DR EMBL; X68383; CAA48453.1; JOINED.
 DR EMBL; BC054758; AAH54758.1; -.
 DR EMBL; BC057931; AAH57931.1; -.
 DR PIR; I48278; KHMSD.
 DR HSSP; P07339; 1LYB.
 DR MEROPS; A01.009; -.
 DR MGD; MGI:88562; Ctsd.
 DR InterPro; IPR001969; Aspprotease_AS.
 DR InterPro; IPR009007; Pept_A_acid.
 DR InterPro; IPR001461; Peptidase_A1.
 DR Pfam; PF00026; asp; 1.
 DR PRINTS; PR00792; PEPSIN.
 DR PROSITE; PS00141; ASP_PROTEASE; 2.
 KW Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Zymogen; Signal.
 FT SIGNAL 1 20 POTENTIAL.
 FT PROPEP 21 64 ACTIVATION PEPTIDE (POTENTIAL).
 FT CHAIN 65 410 CATHEPSIN D.
 FT ACT_SITE 97 97 BY SIMILARITY.
 FT ACT_SITE 293 293 BY SIMILARITY.
 FT DISULFID 91 160 BY SIMILARITY.
 FT DISULFID 110 117 BY SIMILARITY.
 FT DISULFID 284 288 BY SIMILARITY.
 FT DISULFID 327 364 BY SIMILARITY.
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 FT CARBOHYD 261 261 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
 SQ SEQUENCE 410 AA; 44954 MW; DC4928EC46928BF0 CRC64;

Query Match 11.4%; Score 306.5; DB 1; Length 410;
 Best Local Similarity 27.5%; Pred. No. 6.9e-15;
 Matches 103; Conservative 64; Mismatches 123; Indels 85; Gaps 15;

Qy 92 YYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYIDT-----YFDTERSSTYRSKGFDV 145
 || :: ||||| :: ||||| | :| ::::| || |
 Db 79 YYGDIGIGTPPQCFTVVFDGSSNLWVPSIHCKILDIAWVHHKYNSDKSSTYVKNGTSF 138

Qy 146 TVKYTQGSWTGFVGEDLVTIPKGENTSFLVNIAT---IFESENFFLPGI-----KWNGIL 197
 : | || :|:: :| |::| : | | || | || |::|||
 Db 139 DIHYGSGSLSGYLSQDTVSVPCSKSDQSKARGIKVEKQIF-GEATKQPGIVFVAAKFDGIL 197

 Qy 198 GLAYATLAKPSSSLETFFDSLVTQANI-PNVFSMQMCGAGLPVAGSGTNGGSLVLGGIEP 256
 |: | :: :: ||:|: | : |:| | | || |:| || :
 Db 198 GMGYPHIS--VNNVLPVFDNLMQQKLVDKNIFSFY-----LNRDPEGQPGGELMLGGTDS 250

 Qy 257 SLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKAIVD SGTTLLRLPQKVF 316
 | |: | : : |:|: : :||:| : | | |: :|||:| |:| | :
 Db 251 KYYPHGEISYLNVT RKAYWQVHMDQLEVGNE-LTL-CK--GGCEAIVDTGTSLLVGPVEEV 306

 Qy 317 DAVVEAVARASLIPEFSDGFWTGSQ LACWTNSETPWSYFPKIS IYLRDENS SRSFRTIL 376
 : :|: || : | :: |
 Db 307 KELQKAIGAVPLI-----QGEYMIPCEKVSSL 333

 Qy 377 PQLYIQPMMGAGLNYEC----YRFGIS-----PSTNALVIGATVMEG 414
 | :|:: :| | || | :| || :|| :
 Db 334 PTVYLK--LG-GKNYELHPDKYILKVSQGGKTICLSGFMGMDIPPPSGPLWILGDVFIGS 390

 Qy 415 FYVIFDRAQKRVGFA 429
 :| :||| |||||
 Db 391 YYTVFDRDNNRVGFA 405

Search completed: March 4, 2004, 15:36:24
 Job time : 18.5319 secs